

Machine learning modelling for predicting non-domestic buildings energy performance: a model to support deep energy retrofit decision-making

Saleh Seyedzadeh^a, Farzad Pour Rahimian^{b,*}, Stephen Oliver^b, Sergio Rodriguez^b, Ivan Glesk^a

^a*Faculty of Engineering, University of Strathclyde, Glasgow, UK*

^b*School of Computing, Engineering & Digital Technologies, Teesside University, Tees Valley, Middlesbrough, UK*

Abstract

Non-domestic buildings contribute 20% of the UK's annual carbon emissions. A contribution exacerbated by its ageing stock of which only 7% is considered new-build. Consequently, the government has set regulations to decrease the amount of energy take-up by buildings which currently favour deep energy retrofitting analysis for decision-making and demonstrating compliance. Due to the size and complexity of non-domestic buildings, identifying optimal retrofit packages can be very challenging. The need for effective decision-making has led to the wide adoption of artificial intelligence in the retrofit strategy design process. However, the vast retrofit solution space and high time-complexity of energy simulations inhibit artificial intelligence's application. This paper presents an energy performance prediction model for non-domestic buildings supported by machine learning. The aim of the model is to provide a rapid energy performance estimation engine for assisting multi-objective optimisation of non-domestic buildings energy retrofit planning. The study lays out the process of model development from the investigation of requirements and feature extraction to the application on a case study. It employs sensitivity analysis methods to evaluate the effectiveness of the feature set in covering retrofit technologies. The machine learning model which is optimised using

*Email: f.rahimian@tees.ac.uk,

Address: CE2.20, Orion Building, Stephenson St, Tees Valley, Middlesbrough TS1 3BA, UK

advanced evolutionary algorithms provide a robust and reliable tool for building analysts enabling them to meaningfully explore the expanding solution space. The model is evaluated by assessing three thousand retrofit variations of a case study building, achieving a root mean square error of $1.02 \text{ kgCO}_2/\text{m}^2 \times \text{year}$ equal to 1.7% of error.

Keywords: Building energy performance, Data-driven model, Energy performance certificate, Machine learning, Non-domestic building emission rate

Nomenclature

ANN	Artificial Neural Network
BER	Building Emission Rate
DEC	Display Energy Certificate
DHW	Domestic Hot Water
DM	Decision-Making
DT	Decision Trees
EPC	Energy Performance Certificate
EPI	Energy Performance Indicator
EUI	Energy Use Intensity
GA	Genetic Algorithm
GBRT	Gradient Boosting Regression Tree
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MEES	Minimum Energy Efficiency Standards
ML	Machine Learning
MOO	Multi-Objective Optimisation
NDBSC	Non-Domestic Building Services Compliance
R^2	Coefficient of determination
RF	Random Forest
RMSE	Root Mean Squared Error
SBEM	Simplified Building Energy Model
SER	Standard Emission Rate
SVM	Support Vector Machine
\hat{y}	predicted value by ML model
\bar{y}	mean value of the target variable

1. Introduction

In the UK, Non-domestic building stock contributes roughly 20% of annual carbon emission [1]. The UK government has set a series of regulations to abate its gas emission by 29% by 2020 and at least 80% by 2050. A great deal of the problem faced by the UK government and other countries is in existing buildings. In the UK as well as some European countries, the rate of replacing buildings is as low as 1% [2]. It is expected that at least 70% of existing buildings will be occupied in 2050 [3]. However, due to the non-domestic sector's inattentive response to the sustainability program, governments mediation for improvement of energy efficiency for this sector has intensified significantly [4]. Retrofit decision-making, particularly for complex buildings, remains a significant challenge for retrofit analysts [5, 6], given the solution space is effectively unbounded. Decision-making may be further convoluted through legislative or voluntary constraints such as the private rented sector Minimum Energy Efficiency Standards (MEES) seven-year payback exemption criterion [7]. In the case of MEES, retrofit packages which can be demonstrated will not negate initial capital expenditure within seven years can be removed from the compliance solution space.

Therefore, a decision-making (DM) tool is imperative to propose appropriate retrofit technologies for each specific case[8]. Where there is more than one objective, Multi-Objective Optimisation (MOO) methods have become the de facto approach in facilitating decision making and selecting suitable solutions for deep energy retrofits for more than a decade [9]. The majority of existing MOO tools heavily rely on computational energy simulation, where an optimisation script intermittently requests performance estimates from a Building Performance Simulation (BPS) tool such as EnergyPlus [10], TRNSYS [11] and ESP-r [12] to produce energy-related performance indicators. However, despite significantly reducing the time necessary to identify near-optimal retrofit packages, the process and time intensity of such BPSs is highly prohibiting, especially when it comes commercial buildings due to their size and complexity [13, 14].

As such, although AI techniques have been advanced in the last decade, the optimisation of non-domestic, particularly large-scale buildings, faces several challenges towards meeting modern energy standards. These issues are raised from the lack of interdisciplinary researches in this field.

Furthermore, identification and preparation of meaningful and reliable data for modelling have been a primary concern. Due to the high labour and computational costs of energy simulations, neither the retrofit industry nor the stakeholders are willing to use these methods [15]. Thereby, those approaches are limited to academic case studies and partially optimisation of building characteristics [16, 17, 18, 9]. Hence, achieving a comprehensive retrofit planning considering all available technologies and energy policies is not practically possible without a fast and stable energy performance emulator. Otherwise, the means of generating optimal solutions for massive commercial buildings with the development of MOO algorithm will remain exorbitant even for several sample cases.

Motivated by these gaps in the knowledge and practice, this paper concentrates on modelling Building Emission Rate (BER), which is used as the primary index for reflecting non-domestic buildings energy efficiency in the UK. BER is calculated using the Building Regulations UK Part L [19] compliance tool, namely the Simplified Building Energy Model (SBEM) that provides an analysis of a building’s energy performance. The procedure considered available retrofit technologies on the market to cover all alterations in energy performance estimation caused by the application of them. The outcome of the study provides a robust and accurate BER estimation tool to be used in energy efficiency optimisation of complex and heterogeneous buildings. Hence, the model uses a broad dataset of existing buildings rather than notional ones with varying characteristics. Furthermore, the model, coupled with sensitivity analysis allows for identification of the contribution of each input parameter to the overall building performance. Hence, not only the model accelerates the calculation of ratings, but also provides a good insight to recognise the most effective retrofit technologies. Therefore, it will also facilitate the development and implementation of policies and practices regarding buildings energy efficiency. Contributions to carbon footprint reduction will be sought by decreasing the energy consumption, where both community and stockholders will be the beneficiaries of the provided service.

First, a background of building energy benchmarking using data-driven modelling is reviewed. The original building models, available parameters and created synthetic data are elaborated. Then feature extraction and engineering procedures are described in detail. A Machine Learning (ML) model based on a decision tree algorithm is tuned and trained. The model fitting is followed by sensitivity analysis to demonstrate the importance of

the input variables for final selection. Afterwards, the performance of the model in predicting energy performance of a non-domestic building and its retrofitted suggested recommendations (reserved as test cases) is evaluated. Finally, recommendations and discussion on future works are presented.

2. Background

Wang *et al.* [20] described that building energy assessment is an informative tool that provides a comparative energy performance index to decision-makers for energy consumption improvements. Burman *et al.* [21, 22] categorised energy performance assessment regarding engineering methods into top-down and bottom-up approaches. The former entails designing a system with neglect to the information of sub-systems and calculates the incorporated energy rates considering different general building materials. The bottom-up technique involves an accumulation of building system-level details and comparison to actual building efficiency to create a more accurate summary [23].

Borgstein *et al.* [24] addressed a model-based and empirical benchmarking and classified it into three main categories: engineering calculation, simulation model-based benchmarking and statistical modelling. The engineering methodologies employ physical laws for the derivation of building energy consumption in whole or sub-system levels. Building energy efficiency simulation includes software and computer models for simulation of performance with predefined status [25]. The statistical methods use building historical data and frequently apply regression to model the energy consumption/performance of buildings. These models are also called data-driven surrogate models as they take advantage of existing data instead of relying upon intricate system detail [26].

Statistical models are utilised in benchmarking by introducing an anticipated value of energy usage for each building. This method uses different building characteristics as input variables and energy indicator as target values for developing a linear or non-linear model to assess other buildings [26]. The traditional statistical method that has been used widely in the building sector is simple and multivariate regression models [27]. Another popular method is change point regression model, which imitates the non-linear behaviour of input features [28]. Stochastic frontier analysis which is a developed ordinary least squares regression introduces a method

for inefficiency calculation rather than only a simple error measurement [29]. Data envelopment analysis [30] which is a non-parametric method, in contrast with linear regression does not provide any information on the relation of building physical characteristics [31]. By increasing the vast amount of valid and attainable datasets of buildings, there is a great interest in the utilisation of Artificial Intelligent (AI) methods such as ML in the construction sector [26].

2.1. Machine learning for whole building energy estimation

ML is generally used to describe computer algorithms that learn from existing data. These algorithms normally use a large amount of data and a relatively small number of input features for learning processes [32]. ML models operate as a black box and find the relation between the selected input building features and the given energy index, so further information about the building is not required. These features are not necessarily raw building characteristics or weather data; instead, they could be complex variables calculated from basic ones, e.g. wall to floor ratio and mean daily global radiation [33]. The most applied ML techniques in building energy performance assessment are Artificial Neural Network (ANN), Support Vector Machine (SVM), and ensemble models [26].

Yalcintas [34, 35] used ANN for energy benchmarking in tropical climate contemplate weather and chiller data. The selected building includes office, classroom, laboratory-type buildings, or mixed-use buildings. The accuracy of energy use intensity prediction is compared with multiple linear regression method showing a remarkable advantage over it. Wong *et al.* [36] used ANN for assessing the dynamic energy performance of a commercial building with day-lighting in Hong Kong. EnergyPlus software along with algorithms for calculation of interior reflection, was applied to generate the building daily energy usage. ANN can be used for the determination of parameters for energy performance assessment of buildings. Lundin *et al.* [37] proposed a method for prediction of total heat loss coefficient, the total heat capacity and the gain factor that are key elements in the estimation of energy efficiency [32]. Buratti *et al.* [38] employed ANN for evaluation of building energy certificates accuracy using 6500 energy labels in Italy. The study investigated a different combination of input variables to minimise the number of training features. Hong *et al.* [21] applied ANN for benchmarking of school buildings in the UK and investigated the limitations of the assessment. An extensive database, including 120000 records was used for training and testing the

model [39]. Khayatian *et al.* [40] predicted energy performance certificates for residential building using an ANN model and Italian CENED database as training records. A combination set of direct and calculated features was used as inputs to predict heat demand indicators. Ascione *et al.* [41] proposed an ANN for evaluation of energy consumption and inhabitants' thermal comfort to predict the energy performance of new and retrofitted buildings separately. At the same time, Beccali *et al.* [42] proposed the use of ANN fast forecasting as a decision support tool for optimising the refurbishment actions of buildings located in Italy.

In 2010, Li *et al.* [43] used SVM for long-term prediction (yearly) of electricity consumption of domestic buildings. They considered fifteen building envelope parameters collected from 59 different cases along with the annual electricity consumption, which is normalised by unit area. Besides, they compared the accuracy of the SVM model with three types of ANNs, including propagation, RFB and general regression. Testing the trained model over 20% of study cases provided results that showed SVM outperforms ANNs for all samples. Solomon *et al.* [44] predicted weekly electricity consumption of a massive commercial building considering previous electricity usage, temperature data and wind velocity.

The use of ensemble ML models (e.g. RF) in the building energy domain is restricted to only recent years [45, 46, 47, 48, 49], despite a very established track-record of utilisation in other disciplines. Papadopoulos *et al.* [50] also compared different ensemble models in estimation of the energy performance of residential buildings (including 768 variations of a model building) evaluated using Ecotect software. In a recent study [49] which tuned and compared the most commonly used models revealed ensemble models outperforms the alternatives. It was ascertained that the standard and advanced Gradient Boosted Regression Trees (GBRTs) produce the most accurate predictions.

2.2. Energy performance calculation

The aim of building energy performance benchmarking has been to inspire the owners to consume more efficiently [51]. Hence, there has been a necessity for a legal framework to exert the potential of the benchmarking in reducing energy consumption and consequently, carbon emissions. The Energy Performance Certificate (EPC) and Display Energy Certificate (DEC) programs were adopted in the UK to satisfy EU Energy Performance of Building Directive (EPBD) requirements [52]. The EPC

scheme is developed to express the energy-efficiency of buildings regarding their asset ratings which are calculated using simulations. These ratings indicate the carbon emissions of a building as built or as designed. Therefore, EPCs aim at the performance assessment of buildings concerning their fabric and services [53]. In contrast, DEC is a certificate which denotes the efficiency of the operational energy consumption. The principal distinction from the EPC and DEC is whether performance is estimated from simulation or operation. The certificate also includes the inefficient uses of the energy [53]. Hence DEC encourages the occupants to behave in a more energy-efficient manner.

Operational ratings are calculated by comparing the carbon emissions of the building with a representative building. The benchmark indicates the typical performance of similar buildings and can be notional or actual [54].

Operational Rating =

$$\frac{\text{Actual building emission rate (kgCO}_2\text{/m}^2 \times \text{year)}}{\text{Adapted energy benchmark (kgCO}_2\text{/m}^2\text{/year)}} \times 100 \quad (1)$$

Asset ratings for non-domestic EPCs are produced based on the government's Standard Assessment Procedure (SAP) for new buildings. Asset rating is calculated as:

$$\text{EPC Index} = \frac{\text{Actual building emission rate (kgCO}_2\text{/m}^2 \times \text{year)}}{\text{Standard emission rate (kgCO}_2\text{/m}^2\text{/year)}} \times 50 \quad (2)$$

The SBEM tool calculates the energy requirements for the expected levels of heating, hot water, cooling, fans and pumps and fixed internal lighting. These levels are based on the geometry, construction types and assigned activities, along with the services system types, efficiencies and control corrections. Upon calculation of the aggregate energy demand, SBEM transforms the energy into CO_2 emission (i.e. BER).

The Standard Emission Rate (SER) is determined by applying a fixed improvement factor to the emissions from a reference building. EPCs are intended to send market signals about the relative performance of comparable buildings, and so it is necessary that the reference building should be the same for all buildings of a given type. The reference building specifications introduced in 2015 are assigned based upon the servicing strategy identified for each zone within the proposed building. There are therefore two notional

building specifications for a) Heated and naturally ventilated zones, and b) Heated and mechanically ventilated or heated and cooled zones [55].

The ‘reference building’, is a version of the actual building modified in accordance with rules relating to glazing area, insulation and system efficiency. Each space must contain the same activity as proposed for the equivalent space in the actual building. The reference building is generated by SBEM during processing or by accredited software for EPCs modelled using dynamic simulation models such as EnergyPlus. The specification for both developed by the Building Research Establishment (BRE). Both actual and reference buildings have monthly heat balances performed for standard weather data appropriate to the building location. Results are automatically fed into the Building Regulations UK, Part L [19] compliance calculator provided by BRE. The energy performance standards of the reference building is based on a concurrent specification that delivers a 43% reduction in CO₂ emissions relative to the 2010 energy performance standards based on an assumed build mix. This means that the emissions target for some buildings will improve by more than this percentage, others by less [55]. The operational energy performance and asset ratings are displayed in terms of seven-letter ranks, from ‘A’ to ‘G’, where a lower EPC index indicate better performance. It can be seen in equation (2) that a building with an energy consumption similar to the typical performance of buildings in that class will get an operational rating of 50, i.e. between the grades ‘C’ and ‘E.’

This paper concentrates on modelling BER which is used as the main index for reflecting non-domestic buildings energy efficiency in the UK. The reasons for this are that the SER is static and depending on whether the SER is greater than or less than 50 affects the BER/EPC-Index ratio. The outcome of this study is a robust and accurate BER estimation tool to be used in energy efficiency optimisation of complex and heterogeneous buildings. Hence, the model uses a broad dataset of existing buildings rather than representative reference or notional models.

2.3. Decision Making for Retrofit

Building energy efficiency enhancement consists of an optimisation method of determining a selection of technically advantageous and cost-effective measures. The traditional procedure of evaluating a broad variety of retrofit technologies is to investigate several potential solutions based on practician experience. The primary restriction of this strategy is

that only limited number of scenarios can be evaluated and the probability of obtaining an optimal solution is quite low. It has been indicated that by implementing non-optimised solutions, it is possible to alter the building at a subsequent attempt imposing much higher cost [56]. This issue causes investors to be unwilling to invest in their building energy efficiency improvement.

To consider all technologies and combinations, an evaluation of a considerable number of solutions is needed which make the decision-making process a complex work and challenging to manage [57]. Several methodologies have been proposed to facilitated DM in energy retrofits which generally are divided into two groups: a priori or multi-criteria analysis (MCA) and MOO.

MCA still relies on users' experience by defining a set of alternative options and pre-evaluation of the solutions. As such achieving the most optimal retrofit packages is not guaranteed [58]. IN MCA, each criterion is weighted, and then total weights create a unique criterion. Gero *et al.* [59] were among the first to suggest an MCA method for designing energy efficient building. This model examined the trade-off between the thermal efficiency and other factors which are not energy-related (e.g. building cost and available area). This approach was then followed by several researchers by applying in related problems [60, 61, 62, 63]. Jaggs *et al.* [64] and Flourentzou *et al.* [65] suggested strategies for the assessment of retrofitting situations. Kaklauskas *et al.* [66] proposed a multivariate design and MCA approach for deep energy retrofit, defining the importance and advantages of building retrofit options and choosing the topmost preferred alternative. Another drawback MCA based methods is that the information about the sensitivity of each criterion to alteration of others is not provided [67]. Multi-criteria optimisation of CO_2 emission and retrofit cost by considering a few variables including insulation values, air tightness, lighting controls, system efficiencies and PV provision has been reported [13]. It has been pointed out that at least 9,000 simulations are required to obtain optimal solutions. By taking a large number of variables and their options, the number of obligatory simulations will be tremendously increased.

The other method, which is based on MOO, allows considering a broad retrofit technology options limiting the search space and perceive the trade-offs among the objective functions assisting in attaining an optimal solution. Still, comparatively limited consideration has been given to tackling building retrofit DM support with MOO [68]. This is due to the fact that a large

number of simulations are required to obtain a solution. These calculations are time-consuming and demand for heavy processing in case of using energy simulation tools. Diakaki *et al.* [69] studied the usefulness of applying MOO methods for enhancing building energy efficiency, by use of a simplified model for thermal simulation. Asadi *et al.* [68] developed a MOO model to support the definition of retrofit scenarios intended to cost-effectively optimise energy consumption. Further to that, a MOO model coupled with TRNSYS was developed to optimise retrofit cost, energy savings, and thermal comfort in a residential building [70].

Ascione *et al.* [17] proposed a framework for the MOO to optimise building energy design considering the most extensive selection of objective functions and design variables. Genetic Algorithm (GA), combined with EnergyPlus, was used to generate the retrofit solution space for an office building. The same group previously was focused on the use of MOO for optimising residential buildings retrofit planning [71]. Gou *et al.* [72] used the MOO approach by coupling ANN and GA to optimise residential building comfort indoor and energy usage. Ferrara *et al.* [18] also focused on a residential building and used TRNSYS and GenOpt optimisation software to minimise energy demand and global cost. Bre and Fachinotti [73] focused on minimising energy consumptions for heating and cooling and maximising thermal comfort for occupants by coupling EnergyPlus and GA. Jafari and Valentin [74] also employed GA and eQuest and aimed at optimisation of life cycle cost of the retrofit strategies for a residential building. Carlucci *et al.* [75] used GA and EnergyPlus to optimise thermal, visual comfort and indoor air quality in the design stage [76]. With similar configuration and adding ANN, Yu *et al.* [16] aimed at evaluating energy usage and thermal comfort for the design of the Chinese buildings. Pilechiha *et al.* [9] aimed at optimisation of window design in an office room considering the trade-offs among the quality of views, daylight and building energy loads and employing various tools including EnergyPlus, Grasshopper and hypervolume-based evolutionary optimisation algorithm. Another research proposing an EnergyPlus-based optimisation for retrofitting office buildings has reported weeks of simulation for a single case [14]

3. Methodology

Figure 1 illustrates the procedure of research presented in this paper.

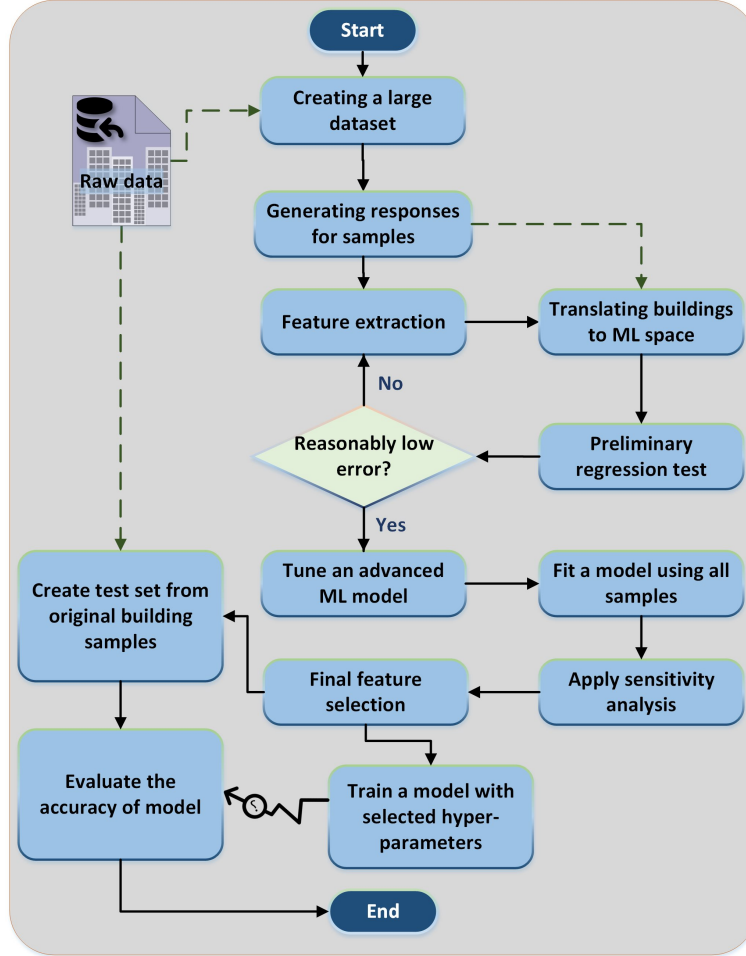


Figure 1: Flow chart of the procedure method.

Before applying a ML, it is necessary to transform the raw building data into a form which is amenable to learning. This procedure includes extracting meaningful input variables (features) from the available data. Each application requires specific considerations when applying this transformation. As the aim of this study was to provide a fast energy efficiency estimation tool for retrofit planning applications, it was essential to take all the available retrofit technologies into account. Hence, the model should produce accurate energy performance predictions, as well as the effect of any alteration due to a building upgrade. The features could be numeric or categorical. However, categorical variables are not preferable as

they are converted to binary variables representing each possible value which increases the feature space and consequently the time complexity.

When the features are selected all data is transferred to the ML space. At this step, training ML models is essential to evaluate the accuracy of predictions using the chosen set of variables. ML requires a reasonable amount of data, and the more precise data leads in a more accurate model.

3.1. Machine learning and gradient boosted regression trees

A critical task in employing ML techniques is the optimisation of model parameters [49]. This procedure is known as tuning, which plays an essential role in the accuracy of ML model predictions. Tuning becomes a more critical task when using complex models such as ANN and GBRT. Picking unsuitable hyperparameters will lead to inaccuracy, which may incorrectly be interpreted as the model failure. Tuning a model is as important as selecting the appropriate input variables for creating and effective ML model. Every ML model has various hyperparameters or meta-parameters which control the learning and prediction processes. A fundamental point in tuning hyperparameters is the generalisation. It means that how well the learning model predicts the specific examples not seen by the model when it was training. Hence, in the procedure of model optimisation, there should be a proper mechanism such as cross-validation to avoid this issue, which is called over-fitting (i.e. representing the training data well but not well suited to unseen data).

3.1.1. Gradient boosted regression trees

GBRT is an ensemble of decision trees (DT) which is a non-parametric model, creates tree-like structures [77]. This ML model continuously splits the data into smaller and smaller subsets and stops when only one record remains in every subgroup. These inner groups and the solitary sets are called nodes and leaves, respectively. However, the accuracy of DT is extensively dependent on the distribution of samples in the training set. Hence, they are not stable models as small changes in the observations will affect the entire structure. One promising solution for that issue has been the utilising a series of DTs and calculating the mean predicted values of all independent trees. The first variation of DT is known as random forest which employs bagging to connect separate models but containing similar information and creates a linear combination from several independent trees. The significant distinction separating GBRT and RF is that the latter one generates fully

exploited DTs, which are all independent of each. Whilst GBRT applies small trees (weak learners) with high bias and low variance [78].

Indeed, these trees are not independent of each other; alternately, every branch is built on the basis of the former simplistic models using a weighting method. This procedure is boosting algorithm in which every inner node splits the given sample set to two branches. Assuming a three nodes GBRT, then there will be one division node in which the best partition of the data is determined, and the deviation of the obtained values are estimated. By training over those residuals, the succeeding DT will explore for a new segment of data to decrease the error variance.

The main hyperparameters for tuning a GBRT model include learning rate that is a weighting method to stop over-fitting by examining the participation of each branch. The number of trees, the maximum depth of the tree and the number of features for searching for the best division, and the minimum number of data samples to split a node and required in each node. Moreover, the sub-sample parameter defines the fraction of observation to be selected for each tree.

3.1.2. Performance evaluation

To assess the accuracy of the trained model, several performance measures about actual and predicted results are calculated. In this work, RMSE, MAE and the coefficient of determination (R^2) were used to present the accuracy of ML models. R^2 is the percentage variance in the dependent variable explained by the independent ones. These values are calculated as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum (y_i - \hat{y})^2} \quad (3)$$

$$MAE = \frac{1}{N} \sum |y_i - \hat{y}| \quad (4)$$

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} \quad (5)$$

Here, y , \hat{y} and \bar{y} represent the real, estimated and average response values, respectively.

Validation is the essential method used for the assessment of the ML model stability and the demonstration of model generalisation. It represents how well the model performs in predicting unseen data (the dataset which is not used to train the model). In this study, cross-validation, a standard

statistical re-sampling method, was employed. In this method, each dataset is randomly divided into k folds, including a training subset, which is used for training the ML model, and a testing subset. These are used to evaluate the model's generalisation efficiency. Then, the average of all k folds accuracies (e.g. RMSE) is calculated and regarded as the final performance. This technique assists the model development procedure in avoiding over-fitting and under-fitting. The former refers to capturing noise and relations which do not generalise accurately to new data. In this case, the trained model runs exceptionally well on the training set, yet poorly at the test set. The latter refers to not capturing relations adequately in the data. In these cases, model accuracy would be poor for both the training and the test sets.

It is conceivable that using all datasets in the cross-validation procedure would not guarantee the performance of the developed model on new data, as the model would be biased to the utilised data. Therefore, it is essential to test the ML model on a set separate from the main one. In this research, the model performance is assessed using an unseen higher education facility which the trained model has never seen in any form. The building detail is processed, and variations are generated via the GA algorithm. The trained model accuracy is then evaluated by predicting those samples. The consistency between the simulation results and the predictions of the proposed model represented in the low RMSE serve as an indicator of external validity for this study

3.1.3. Model parameter optimisation

The importance of ML hyperparameter optimisation forecasting building energy has been highlighted in the literature [49]. The traditional method for tuning ML models has been the exploration of all possible configurations of model parameters using grid search. However, this method can be extremely laborious for complex models as ANN and GBRT for they have many parameters to be tuned. Seyedzadeh *et al.* [79] used a multi-objective optimisation based on Non-dominated Sorting Genetic Algorithm II (NDSGA-II) [80] to tune the model for accurate estimation of cooling and heating loads at the same time. In this study, the model has only one target that is BER, and GA itself needs an alignment for getting optimal results. Hence, this research used Sequential Model-based Algorithm Configuration (SMAC) [81] for smart optimisation of the GBRT models.

Figure 2 illustrates the mechanism of tuning algorithm coupled with cross-

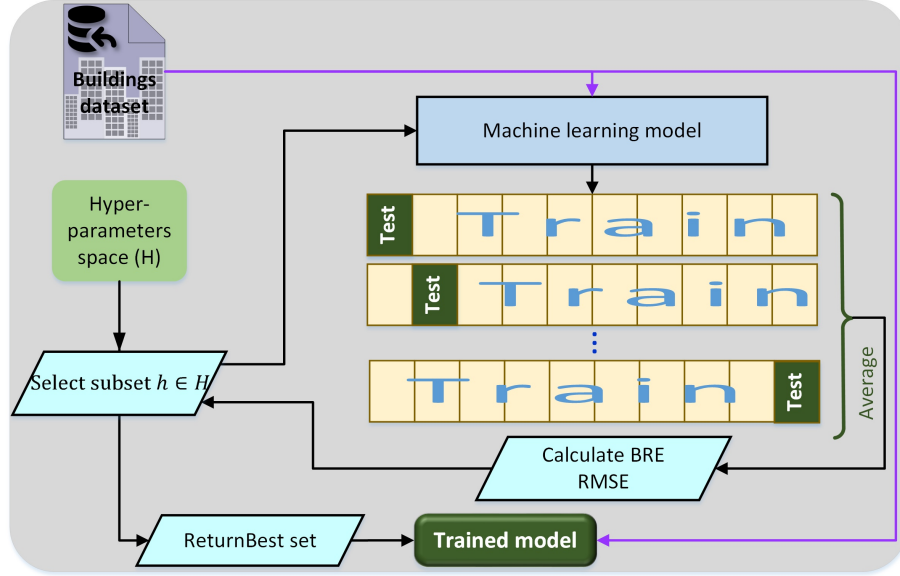


Figure 2: Diagram of the hyperparameter optimisation method for modelling buildings BER.

validation testing. First, the domains of all hyperparameters to be optimised are defined. Next, the algorithm starts with the specified default values and sends it to the evaluation function to build a model for BER prediction. It then continues based on the evaluated RMSE and smartly creates new configurations. Finally, it stops when the number of iterations reaches the maximum value, which, i.e. 800 in this case. Whereas, using the grid search method, the number of repetitions would be around 3,500. As mentioned in the previous section, smart ML optimisation not only reduces the time complexity but also allows for the selection of precise tuning.

3.2. Feature engineering

Feature engineering, as a fundamental element of the application of machine learning, is defined as “*using domain knowledge of the data to create features that make machine learning algorithms work*” [82]. For high-performance prediction by ML, the most explicit and profound relationship of data should be exposed in the form of features. Building energy performance modelling requires strong expertise in building physics and retrofit technologies as well as creative feature extraction methods to interpret this knowledge into useful features.

Feature engineering is not a one-step procedure, but several processes are iteratively applied until a satisfactory result is achieved. Figure 3 demonstrates the life cycle of feature engineering which is adopted in this study.

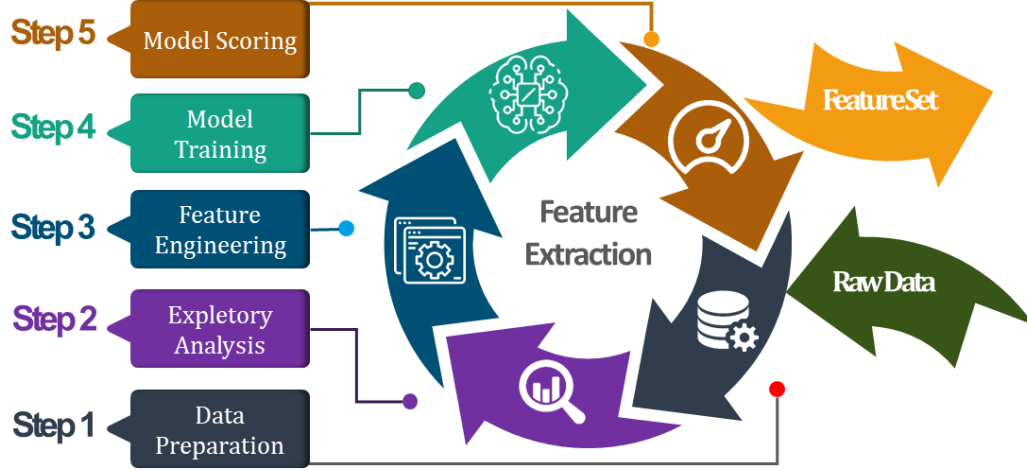


Figure 3: Feature extraction circular procedure.

Firstly, the raw data, including several thousands of non-domestic records, are collected. Potential features related to building characteristics are then selected. This also includes combining datasets from different building energy data to generate unique and independent features. Statistical analysis is then performed to evaluate the impact of derived features. This step involves an investigation of data to identify potential outliers and correlation analysis. The subsequent stage, which is crucial to achieving the main objective of this study, is engineering new features. In this step, it is essential to consider retrofit technologies to be covered by the variables, as the aim of this study was not only to predict energy performance accurately but also to effectively support retrofit decision-making. This process is therefore for creating new features from existing variables based on knowledge on retrofit planning and energy performance calculation (SBEM). It also includes translating categorical variables to the ML model usable form. When the final feature set is decided, a model is trained using the set and then scored using the

evaluation method described earlier in this section. All these steps are repeated until the model performance is satisfactory.

3.2.1. Data preparation

The studied data was obtained from arbnco company’s arbnco Latd’s arbn Consult platform [83]. At the time of this research’s conception, the dataset consisted of 4,900 records representing non-domestic buildings distributed all over the UK. All records were assessed and labelled using the latest version of SBEM software (v5.6.a) considering related existing building regulations. Each building’s model is submitted as “.inp” files by assessors, engineers or owners. The data then went through a full SBEM energy simulation, and the BER extracted from the results. The SER was also retrieved from the results for the reference building. The EPC is calculated based on these rates using Eq. 2. This study aimed at developing an accurate and fast prediction model for building BER.

From these records, around 80,000 samples were mutated using possible alterations to create a large dataset to train a robust and general model. Most of these mutated buildings have better energy performance than their original building states. The method is advantageous for this study, as it aims at creating a model most suitable for predicting energy efficiency of potential retrofitted building states to be used in the optimisation process. All mutated buildings were then evaluated using the software and recorded with their assessed BER values.

3.2.2. Feature selection

Extracting features for modelling building energy indices is dependent on the several parameters including the prediction period (hourly, monthly, annually, etc.), target (electricity, cooling, heating or whole building demand) and the studied case (i.e. investigating one case or a group of records). Short-term predictions usually aimed at modelling the single cases and involved few features mostly weather-related and in some instances the calendar nominal attributes [26]. However, when dealing with the whole building energy performance simulations, a high number of variables should be considered. In this paper, as the focus is to aid retrofit optimisation, it is essential to ensure that every possible alteration will be reflected by one or more input variables. Therefore, it is essential to understand the retrofit solution space before defining the ML input space.

Table 1 presents the considered retrofit options for existing non-domestic buildings in the UK.

Table 1: Available retrofit recommendations for EPC rating improvement

Code	Description
C2	Modify the seasonal and nominal efficiency properties of comfort cooling plant to mimic the replacement of existing chiller with a system compliant with the building regulations as defined in non-domestic building services compliance, guide 2013 (NDBSC-2013), Section 9.
E2	Modify the roof thermal properties to mimic the effect of insulating the roof in accordance with Approved Document L2B: Conservation of fuel and power in existing buildings other than dwellings. Section 5, Retaining thermal elements
E4	Modify envelope thermal properties of cavity walls to mimic the effect of insulating the cavity in accordance with Approved Document L2B: Conservation of fuel and power in existing buildings other than dwellings. Section 5, Retaining thermal elements
E8	Modify the glazing thermal properties to mimic the effect of replacing existing glazing in accordance with Approved Document L2B: Conservation of fuel and power in existing buildings other than dwellings. Section 4, Work on controlled fittings and services
H1	Modify low-temperature hot water boilers' system seasonal energy efficiency to replicate the replacement of inefficient buildings as per NDBSC-2013 Section 2 Gas, oil and biomass-fired boilers
H7	Modify the heating system to mimic the installation of an optimum start/stop controller to the heating system as per the heating efficiency credits system outlined in NDBSC-2013, Section 3: Gas, oil and biomass-fired boilers
H8	Modify the heating system to mimic the installation of weather compensation systems to the heating system as per the heating efficiency credits system outlined in NDBSC-2013, Section 3: Gas, oil and biomass-fired boilers
L2	Modify zonal lighting system properties for zones with Tungsten or Halogen filament lamps to mimic a like-for-like replacement with high-efficiency lamps in accordance with NDBSC-2013, Section 12: Lighting

Table 1 (cont.) Available retrofit recommendations for EPC rating improvement.

Code	Description
L5	Modify zonal lighting system properties for zones with T8 fluorescent tube light to mimic a like-for-like replacement with high-efficiency T5 fluorescent tubes in accordance with NDBSC-2013, Section 12: Lighting
L8	Modify zonal lighting system properties for zones with T12 fluorescent tube light to mimic a like-for-like replacement with high-efficiency T5 fluorescent tubes in accordance with NDBSC-2013, Section 12: Lighting
R5	Modify or create a HVAC whose heating and comfort cooling demand are served by an air source heat pump in accordance with NDBSC-2013, Heat pumps
V1	Modify the g-value of glazing systems to mimic the application of solar control film to all transparent surfaces
W1	Modify the existing domestic hot water system to mimic the installation of a high-efficiency version in accordance with NDBSC-2013, Section 8: Domestic hot water
W2	Modify the centralised domestic hot water systems to mimic replacing the systems with instantaneous point of use systems.
W3	Replace or install domestic hot water cylinder jacket in accordance with NDBSC-2013, Section 8: Domestic hot water
MTR	Improve the building metadata to reflect the nominal energy consumption benefits of installing sub-metering
AFM	Modify existing air handling unit efficiencies to mimic the installation of a high-efficiency system in accordance with NDBSC-2013, Section 10: Air distribution
AHR	Modify the existing mechanical ventilation system to mimic the installation of a high-efficiency rotary heat exchanger in accordance with NDBSC-2013, Section 10: Air distribution
DSF	Modify the properties of zones with high ceilings to mimic the installation of destratification fans.
FC	Modify the energy efficiency properties of existing fan-coil units to mimic the installation of high-efficiency units in accordance with NDBSC-2013, Section 1.7: Summary of recommendation minimum energy efficiency standards

Table 1 (cont.) Available retrofit recommendations for EPC rating improvement.

Code	Description
RAD	Modify existing convection heating systems in industrial and warehouse areas with a ceiling-mount radiant heating system in accordance with NDBSC-2013 Section 5: Gas and oil-fired radiant heaters.
VSD	Modify existing low-temperature hot water boilers to mimic the installation of variable speed pumps.
WET	Modify existing direct electric heating systems with a high-efficiency wet radiator system served by a low-temperature hot water boiler in accordance with NDBSC-2013, Section 2: Gas, oil and biomass-fired boilers
8LO	Modify zonal lighting system properties for zones with T8 fluorescent tube light to mimic a lamp-luminaire replacement with high-efficiency T5 fluorescent tube systems in accordance with NDBSC-2013, Section 12: Lighting
DLD	Mimic the installation of photoelectric daylight dimming controls
HLU	Modify zonal lighting system properties for zones with Tungsten or Halogen filament lamps to mimic a lamp-luminaire replacement with high-efficiency systems in accordance with NDBSC-2013, Section 12: Lighting
HPL	Modify zonal lighting system properties for zones with High Pressure Sodium or High-Pressure Mercury to mimic a lamp-luminaire replacement with high-efficiency systems in accordance with NDBSC-2013, Section 12: Lighting
HPT	Modify zonal lighting system properties for zones with T5 fluorescent tubes to mimic a lamp-luminaire replacement with high-efficiency systems in accordance with NDBSC-2013, Section 12: Lighting
PIR	Modify zonal lighting system properties to mimic the installation of passive infrared occupancy sensors.
T8L	Modify zonal lighting system properties for zones with T8 fluorescent tubes to mimic a lamp-luminaire replacement with high-efficiency systems in accordance with NDBSC-2013, Section 12: Lighting

3.3. Feature importance

A potent tool for identifying the important features is sensitivity analysis. Including ML feature importance and permutation importance, the use of sensitivity analysis in this study will be explained in this section.

Several methods are proposed for the aim of evaluating feature importance [84] including correlation matrices, sensitivity analysis and ML-based methods. The correlation matrix is typically presented as a heat map of Pearson correlation [85] values, and it calculates the linear relationships between each pair of features and range from 1 to +1 (higher absolute values indicating greater interdependence). In this research, to check the significance of the selected variables in the regression model, permutation importance and the trained GBRT model itself is used to determine what features are of the highest importance.

In GBRT models, the more a variable is utilised to make critical determinations with the trees, the greater its relative importance. This value is computed for each feature, enabling them to be ranked. The relative importance is determined for a single tree considering the amount by which each split point enhances the accuracy measure of the model. The average value of all importances for an individual feature in all independent trees is calculated.

For a single decision tree T , the importance for each variable X_l is calculated as

$$\mathfrak{I}_l^2(T) = \sum_{t=1}^{J-1} \hat{i}_t^2 I(v(t) = l) \quad (6)$$

where \hat{i}_t^2 is the maximal estimated improvement. The sum is calculated over $J - 1$ internal nodes at each tree, and one of the features $X_{v(t)}$ is utilised to split the region related to that node into two branches. The single variable selected is the one that provides the most significant expected improvement of \hat{I}_t^2 in squared error risk over that for a regular fit over the whole branch. The average for all the additive M trees is defined as:

$$\mathfrak{I}_l^2 = \frac{1}{M} \sum_{m=1}^M \mathfrak{I}_l^2(T_m) \quad (7)$$

To perform the permutation importance, a fully trained model is also required, but a separate test-set is utilised to evaluate the significance of features. In this method, each input variable is iteratively permuted, while

keeping others constant. This process is repeated until all features are investigated. The model accuracy alteration due to shuffling of each variable is considered as its importance. Although this technique is computationally expensive, it is a useful complement of the ML importance method, especially in identifying inflated values.

Moreover, the sensitivity analysis provides beneficial information on how the model predicts, making it less of a black box. In other words, it reveals which variables play a more significant role in calculating the building energy performance. Hence, this detail can be verified by comparison to the field of knowledge. The influence of the retrofit strategies on the model can also be conceded and substantiated with this method.

4. Result and discussion

4.1. *Extracted features*

The procedure of selecting the features for modelling energy performance was conducted by iterative processes. The primitive models for evaluation of generated input set were trained and tested. At each iteration, new features were defined and the function for extracting them from building “.inp” files were formulated. Some variables with little influence on predictions were removed and several others modified based on a priori knowledge of SBEM and building physics.

Table 2 shows the final extracted features from buildings characteristics available in the logged assessment files and their description. Most of these variables are calculated, though several are directly taken from the input models. Column “Calculation” in Table 2 explains the formulation of those derived features. “Mutable” column determines if a feature is changeable, and if it is, which recommendations affect the variable are shown in the last column, “Covered solutions”. Equivalent is prepended to several features which indicates that the value has been adjusted to account for the difference in energy required of each fuel to produce a kilogram of CO_2 . The adjustment is based on the ratio of the grid-supplied electricity $kgCO_2/m^2$ emissions factor to that of a given fuel. The label applies to features which affect heating/cooling demand.

Table 2: Features extracted for training BER prediction model

Feature	Description	Calculation	Mutable	Covered solutions
AIR	Air Infiltration Rate	Average air leakage rate measured in $m^3/m^2/h$ at 50 Pascal	Yes	H1, H7, H8, R5, MTR, RAD, WET, DSF
CL_E	Peak solar irradiance for East facing surface weighted by the sum of external surfaces of zones with cooling		No	-
CL_N	Peak solar irradiance for North facing surface weighted by the sum of external surfaces of zones with cooling		No	-
CL_NE	Peak solar irradiance for NE facing surface weighted by the sum of external surfaces of zones with cooling		No	-
CL_NW	Peak solar irradiance for NW facing surface weighted by the sum of external surfaces of zones with cooling		No	-
CL_S	Peak solar irradiance for South facing surface weighted by the sum of external surfaces of zones with cooling		No	-
CL_SE	Peak solar irradiance for SE facing surface weighted by the sum of external surfaces of zones with cooling		No	-
CL_SW	Peak solar irradiance for SW facing surface weighted by the sum of external surfaces of zones with cooling		No	-
CL_W	Peak solar irradiance for West facing surface weighted by the sum of external surfaces of zones with cooling		No	-
DHW_Dem	Domestic hot water demand	DHW system demand normalised by area	Yes	W1, W2, W3, VSD
ECS	Seasonal energy efficiency ratio	Area-weighted season efficiency of all chillers serving at least one space	Yes	C2
EHS	Electric equivalent seasonal coefficient of performance	Weighted heating area adjusted by fuel emissions factor	Yes	VSD, R5, WET, RAD

Table 2 (cont.) Features extracted for training BER prediction model

Feature	Description	Calculation	Mutable	Covered solutions
I.E	Peak solar irradiance for East facing external surface weighted by the sum of external surfaces of zones with heating		No	-
I.H	Peak solar irradiance on roof weighted by the sum of external surfaces of zones with heating		No	-
I.N	Peak solar irradiance for North facing external surface weighted by the sum of external surfaces of zones with heating		No	-
I.NE	Peak solar irradiance for NE facing external surface weighted by the sum of external surfaces of zones with heating		No	-
I.NW	Peak solar irradiance for NW facing external surface weighted by the sum of external surfaces of zones with heating		No	-
I.S	Peak solar irradiance for South facing external surface weighted by the sum of external surfaces of zones with heating		No	-
I.SE	Peak solar irradiance for SE facing external surface weighted by the sum of external surfaces of zones with heating		No	-
I.SW	Peak solar irradiance for SW facing external surface weighted by the sum of external surfaces of zones with heating		No	-
I.VC	Solar irradiance, vertical surfaces in cooled spaces (direct +diffuse solar gains peak July)		Yes	E2

Table 2 (cont.) Features extracted for training BER prediction model

Feature	Description	Calculation	Mutable	Covered solutions
I_VH	Solar irradiance, horizontal surfaces in heated spaces (Direct +diffuse solar)		Yes	E2
I_W	Peak solar irradiance for West facing external surface weighted by the sum of external surfaces of zones with heating		No	-
ICO	Heat transfer from conditioned to unconditioned spaces		No	-
IGC	Internal gain from equipment for cooling spaces		No	-
IHG	Internal gains from spaces with comfort cooling	Sum of sensible gains from equipment and lighting, and latent gains from occupants in W	Yes	8LO, DLD, HLU, HPL, HPT, PIR, T8L, L2, L5, L8, E2, E4, E8, H1, H7, H8, R5, MTR, RAD, WET
LPD	Lighting power density (Area-weighted average lighting power density in W/m^2)		Yes	8LO, DLD, HLU, HPL, HPT, PIR, T8L, L2, L5, L8
PL_H	Peak load of heating demand	Heating system design peak load (kW)	Yes	V1

Table 2 (cont.) Features extracted for training BER prediction model

Feature	Description	Calculation	Mutable	Covered solutions
RWR	Roof to wall ratio	Ratio of total roof area to total wall area	No	-
SER	Standard emission rate	Calculated from initial SBEM run for the reference building	No	-
SFP_Vent	Specific fan power, ventilation, terminal units and exhaust (terminal unit energy demand)	One value per HVAC system	Yes	FC, AFM, AHR
SR_C	Cooled space ratio	Ratio of total cooled space to total space	No	-
SR_H	Heated space ratio	Ratio of total heated space to total space	No	-
SR_UC	Unconditioned space ratio	Ratio of total unconditioned space to total space	No	-
SR_Vent	Ventilated space ratio	Ratio of total ventilated space to total space	No	-
ST_ExW	Equivalent solar transmittivity of the external windows	(window) area-weighted average of the solar transmittivity of individual external walls	Yes	E8, V1

Table 2 (cont.) Features extracted for training BER prediction model

Feature	Description	Calculation	Mutable	Covered solutions
U_Opaq	Equivalent U-value of the opaque part of the external surfaces (walls)	(wall) area-weighted average of the U-values of individual external walls	Yes	H1, H7, H8, R5, MTR, RAD, WET
U_Trans	Equivalent U-value of the transparent part of the external surface (windows)	(window) area-weighted average of the U-values of individual external windows	Yes	E8
VSR_adia	Internal adiabatic vertical surfaces ratio	Ratio of total internal adiabatic vertical surface area to total area of all vertical surfaces	No	-
VSR_diab	Internal diabatic vertical surfaces ratio	Ratio of total internal diabatic vertical surface area to total area of all vertical surfaces	No	-
VSR_ext	External vertical surfaces ratio	Ratio of total external vertical surface area to total area of all vertical surfaces	No	-
WFR	Wall to floor ratio	Ratio of total walls area to total floor area	No	-
WWR	Window to wall ratio	Ratio of total window area to total wall area	No	-

Figure 4 illustrates the frequency of features as histogram graphs. The correlation between each pair of input and target variables is demonstrated using heatmap matrix in Figure 5.

As seen in Figure 5, there are no direct correlations between selected variables and BER. High inter-correlations of features result in bias within tree-building for predictions where decision tree models are sensitive to the inter-correlation among the input variables, and forecast would be problematic based on highly correlated inputs [86]. It is not also possible to identify which variable can be removed without affecting the model precision.

4.2. Model optimisation

An ML model was optimised using the generated as well as actual data. The GBRT model was tuned using SMAC algorithm and 5,000 randomly selected record. This research used a five-fold-cross validation for evaluating the performance of each model configuration. The tuning algorithm output after 1,000 iterations was the hyper-parameter set with which the model reached the RMSE accuracy of $7.01 \text{ kgCO}_2/\text{m}^2 \times \text{year}$ (mean RMSE of all folds). Next, the same configuration was used for modelling variant number of records up to 80,000. Each model was tested using ten-fold cross-validation, and the results recorded as the worst, average and best RMSE of all folds, as demonstrated in Figure 6.

As shown in Figure 6, 30,000 samples of building assessment records are adequate to build a reliable model, as at this point, the prediction interval of $[1.57, 2.55]$ with an average RMSE of $1.92 \text{ kgCO}_2/\text{m}^2 \times \text{year}$ is achieved. Considering the average actual BER of $94.04 \text{ kgCO}_2/\text{m}^2 \times \text{year}$ for all building records, the attained error, which is only 2% of the target mean value is a promising achievement. Using the full dataset, the average RMSE of $1.25 \text{ kgCO}_2/\text{m}^2 \times \text{year}$ equal to 1.3% of the target mean value could be achieved with the acceptable cost of sacrificing time. The average spent time for fitting models with 30,000 and 80,000 samples was recorded as 7.12 and 31.22 seconds.

4.3. Sensitivity analysis

While this research is effectively the creation of an energy performance emulator for existing non-domestic buildings, this study is focused on retrofit decision-making. Therefore feature engineering targeted not only an accurate prediction of building energy performance, but also the ability of the model to

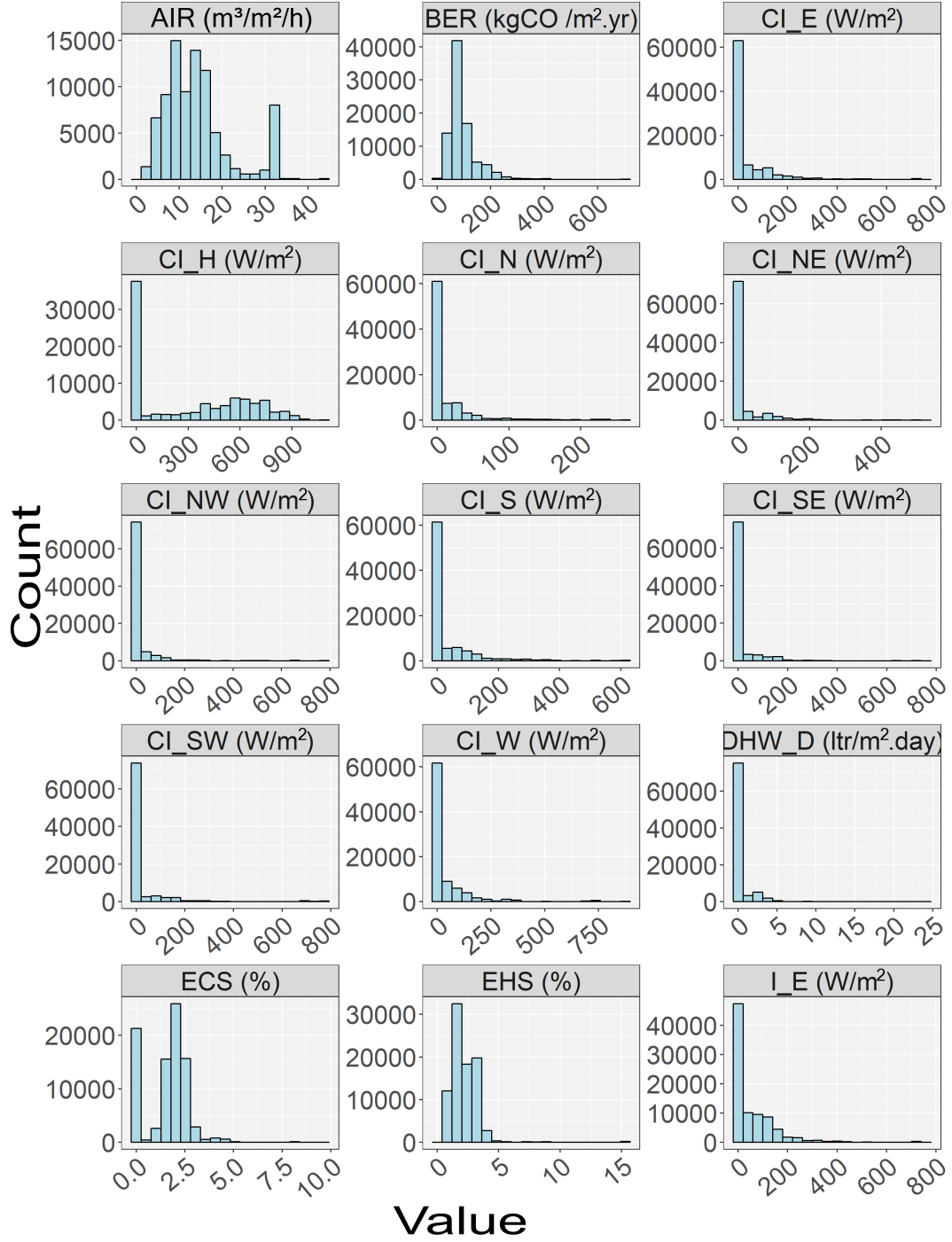


Figure 4: Distribution of the selected features for building energy data.

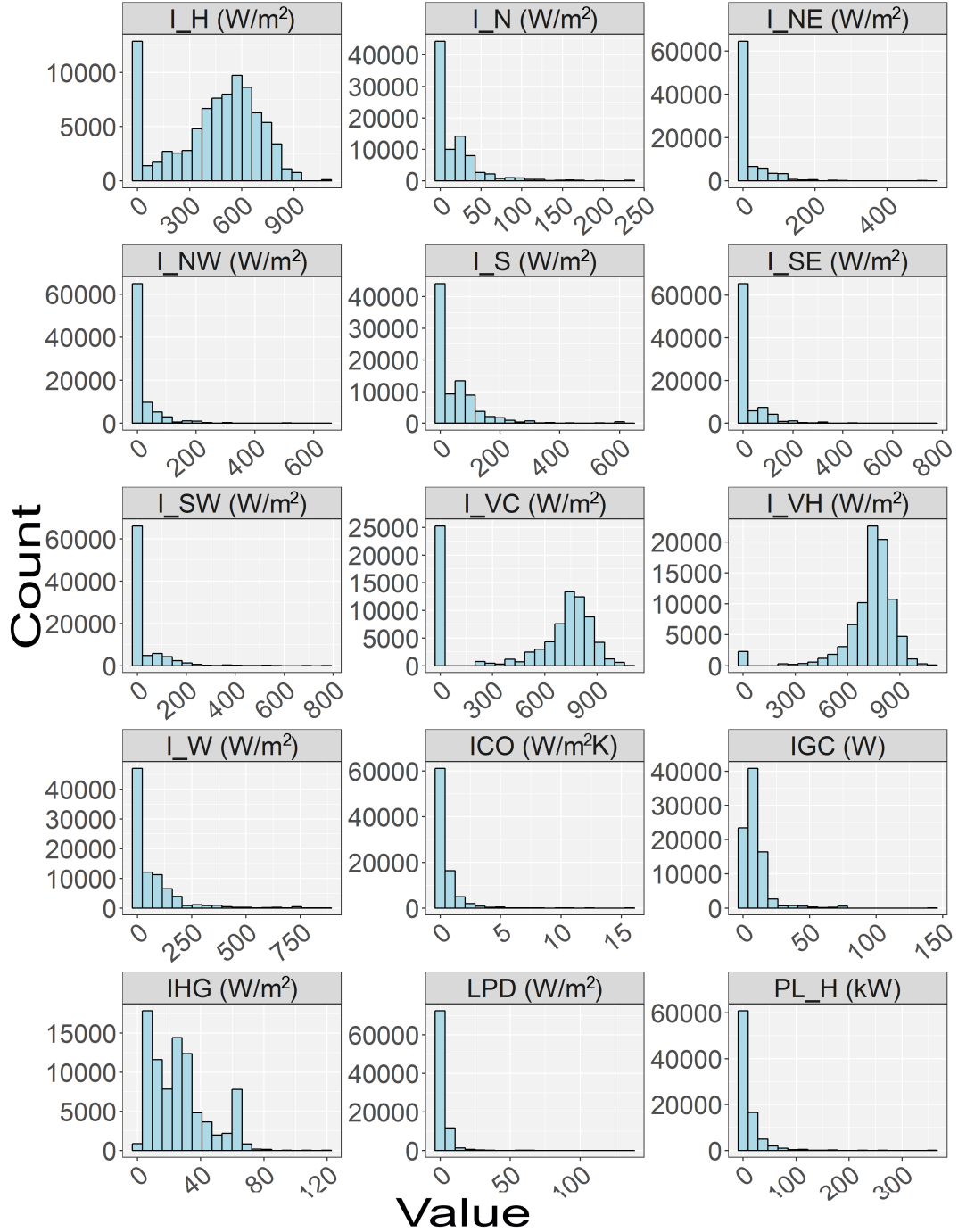


Figure 4 (Cont.): Distribution of the selected features for building energy data.

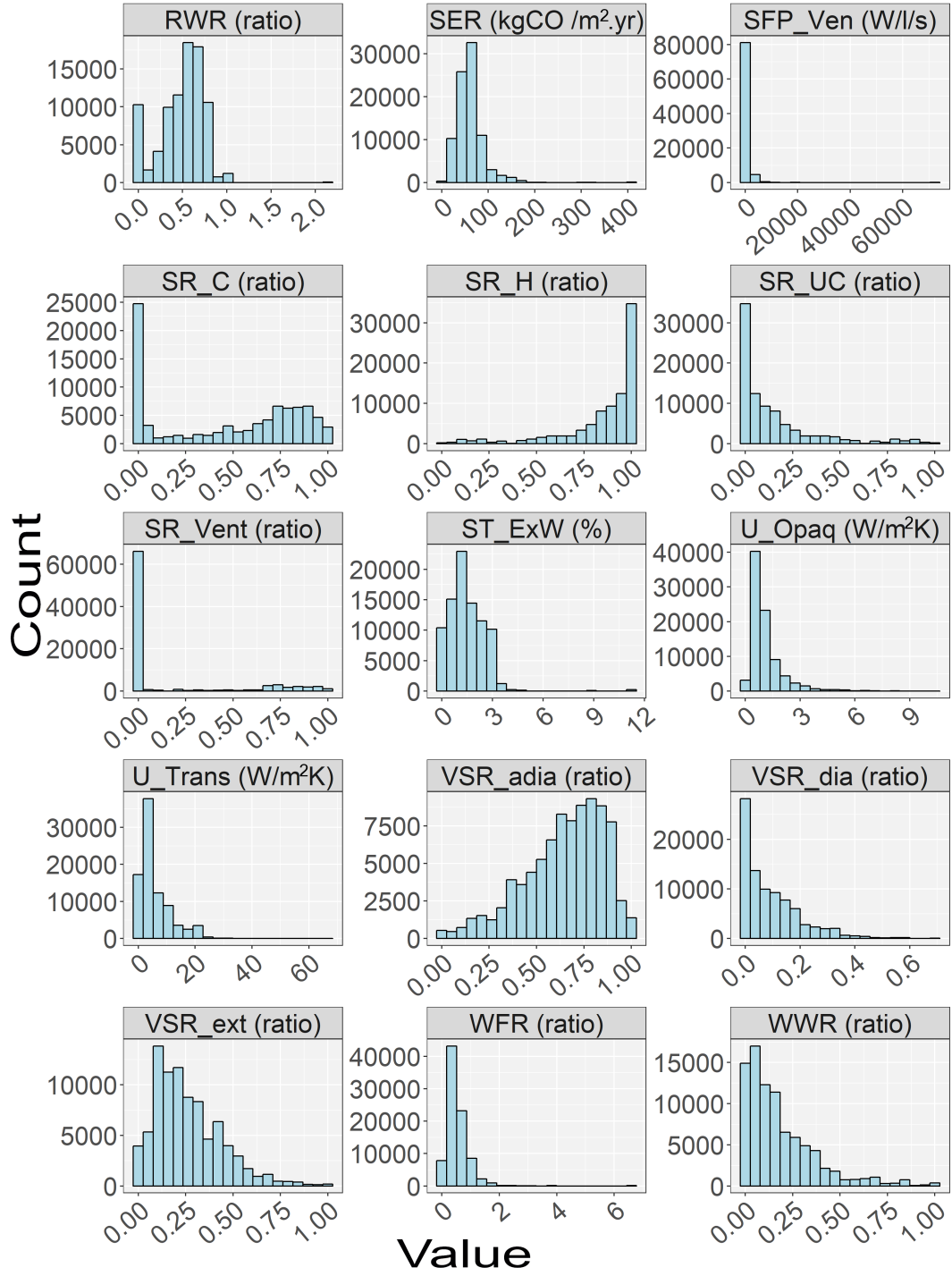


Figure 4 (Cont.): Distribution of the selected features for building energy data.

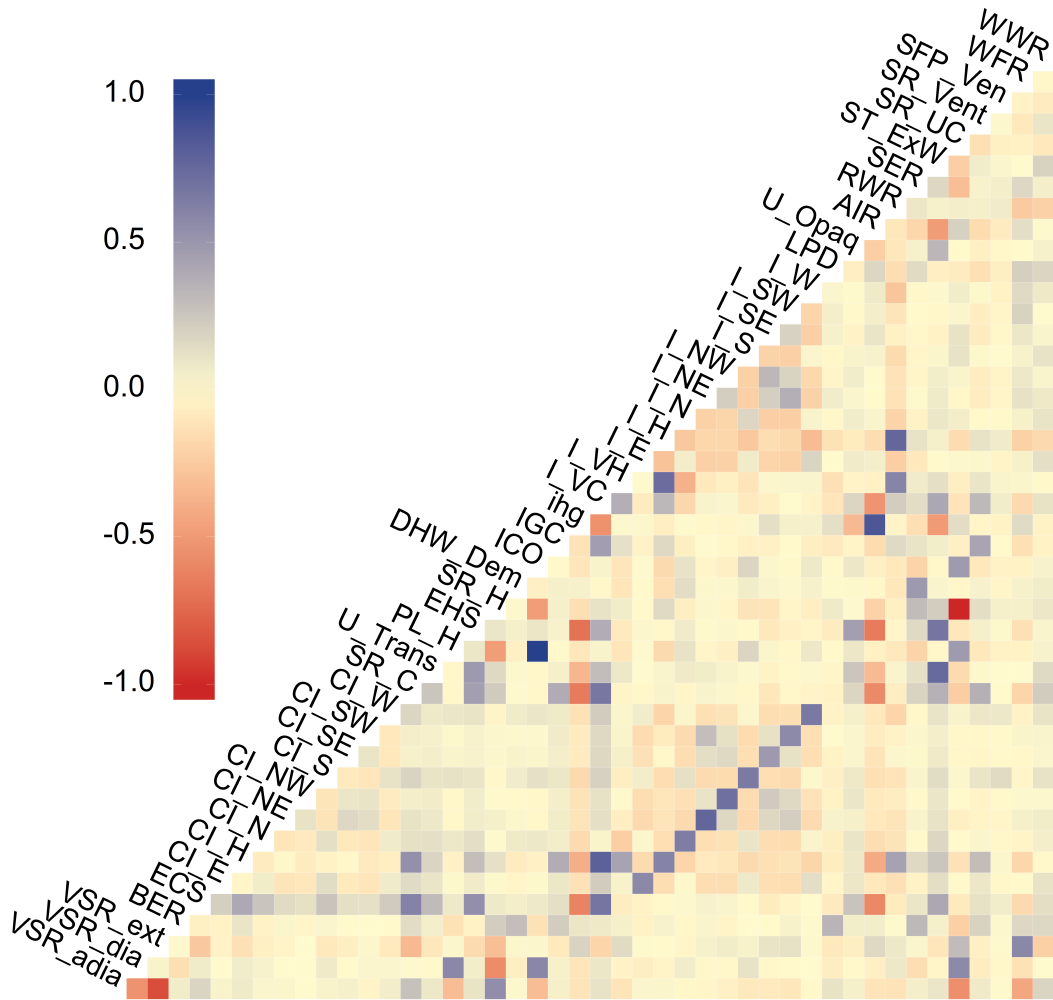


Figure 5: Non-domestic building data features values represented as a heat correlation map.

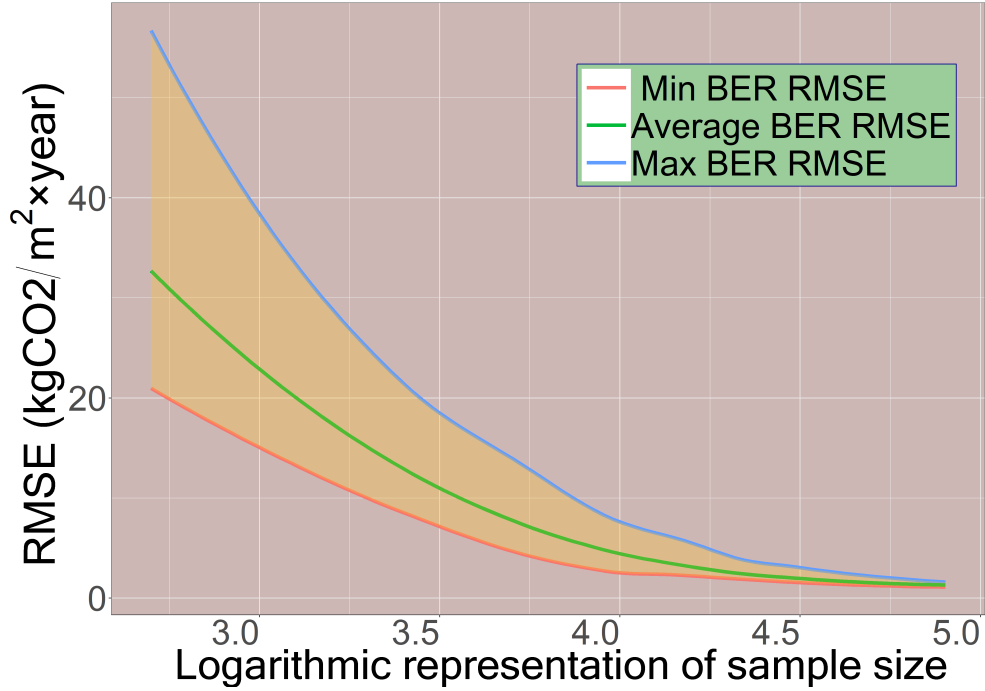


Figure 6: Average, min and max RMSE of all folds in BER prediction against total number of train-test records

predict post-retrofit performance. As such, it is beneficial to investigate the impact of each selected variable on estimating buildings energy performance. To this end, the GBRT model was trained 30 times each using randomly selected records (30,000) and different random states. The results illustrated in Figure 7 plot the relative importance of features as a box and whiskers plot.

The result of permutation importance analysis considering a model trained over 80% of full data and the rest for testing is presented in Figure 8. It can be seen that very similar results to those presented in Figure 7 were achieved. This shows that the GBRT model has not been biased in the training procedure and inflating the feature importances.

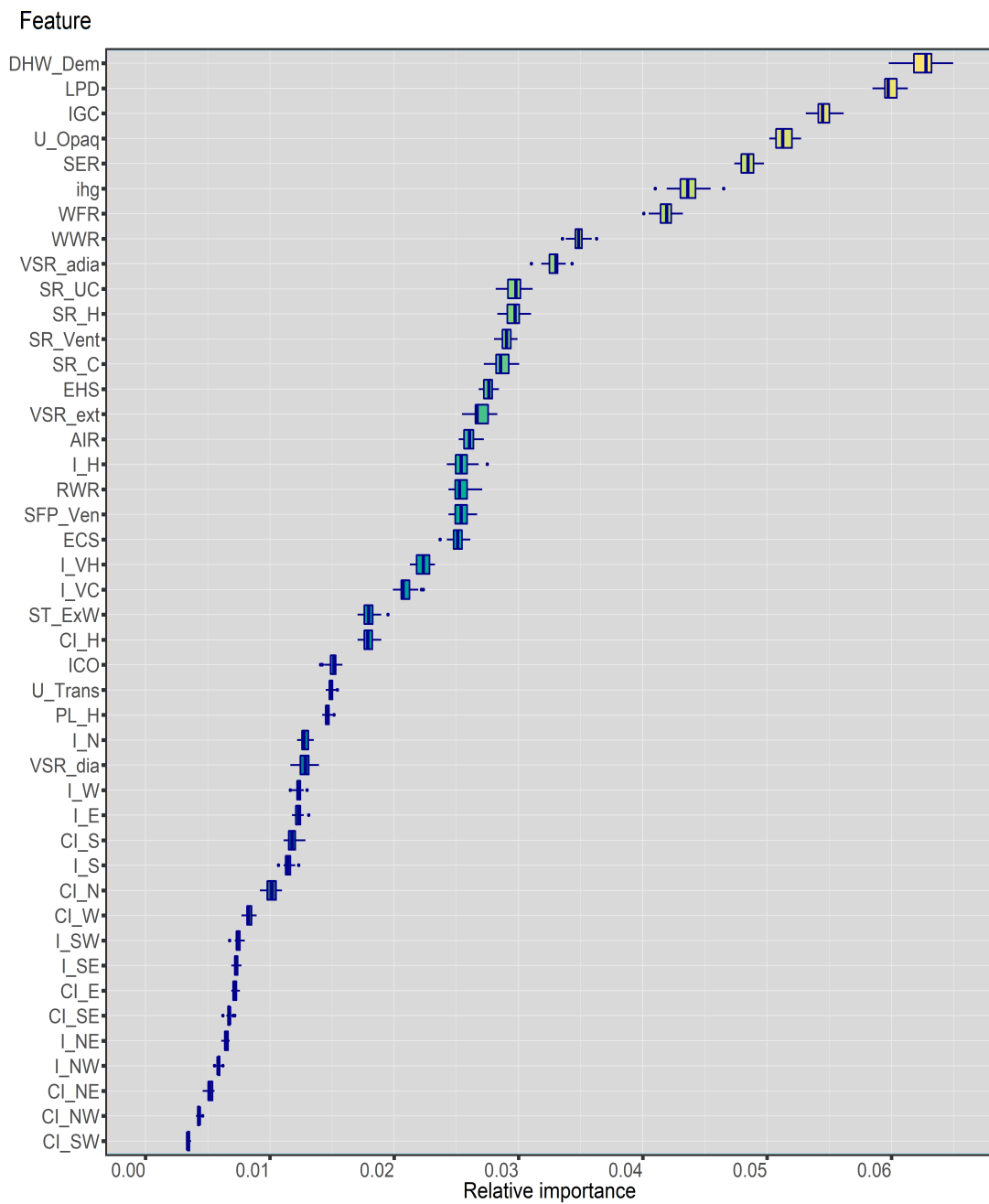


Figure 7: Relative importance of the selected features for modelling non-domestic buildings in the UK using the GBRT model

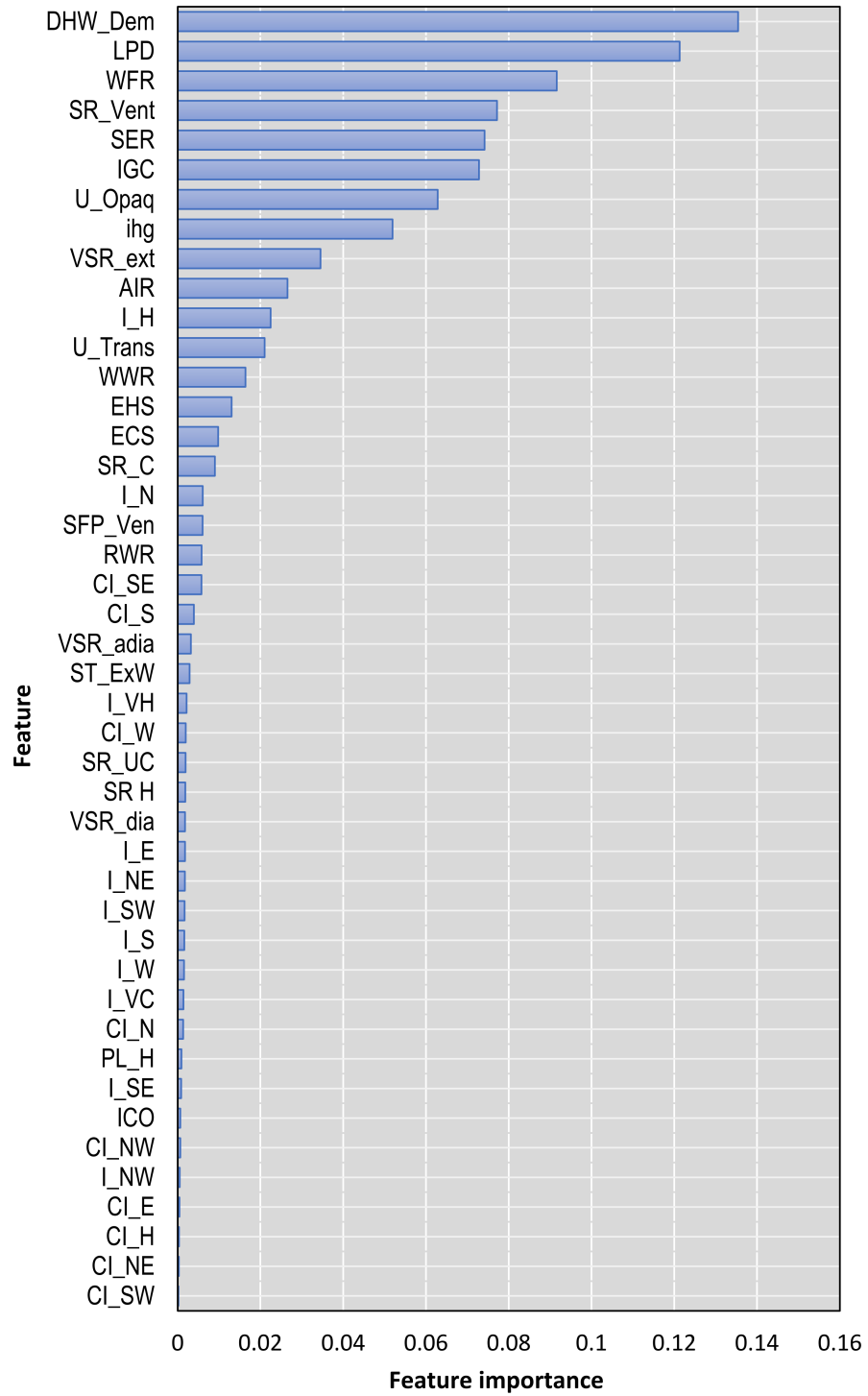


Figure 8: Importance of the selected features evaluated by permutation importance method

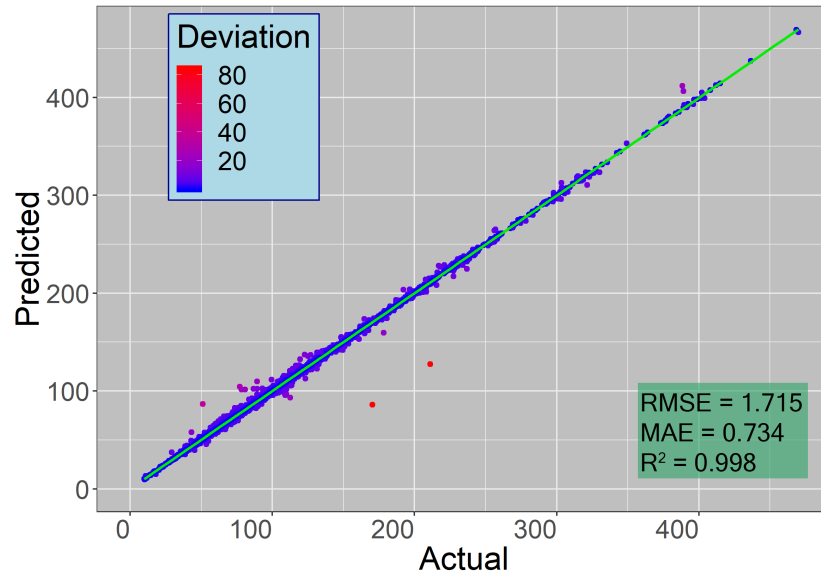
Consequently, the most significant features are related to domestic hot water, internal heat gain and lighting. These parameters cover the improvement of the hot water system, replacing the lamps and heating systems. SER which is indirectly calculated from the building energy assessment procedure, along with the wall to floor and window to wall ratios are not mutable. That is to say, in energy retrofitting the structure of the building is not altered. However, they play an influential role in predicting energy performance. Other immutable features with high impact having almost the same relative importance, are space ratios, followed by air infiltration rate, terminal unit energy demand and cooling system efficiency. As it could be anticipated, the roof to wall ratio, solar radiation on the roof and generally solar radiation have medium to low impact, due to the UK climate. Although the importance of some features of a building is obvious in the field study, the sensitivity analysis here further cross-checks the results of the developed ML model with field results. Conforming this in the model acts as an indicator of external validity of it, and as a means for triangulation.

It should be noted that the importance of the features in the prediction of building energy performance is dependent on the data utilised for developing the model. As such, this conclusion is not generalised for countries with different climates and policies (software to check energy performance). That said, the proposed feature extraction methods may underpin other building energy retrofit decision-making tools and policies. Furthermore, the data and results can be transferred to the new domain for expedite model development.

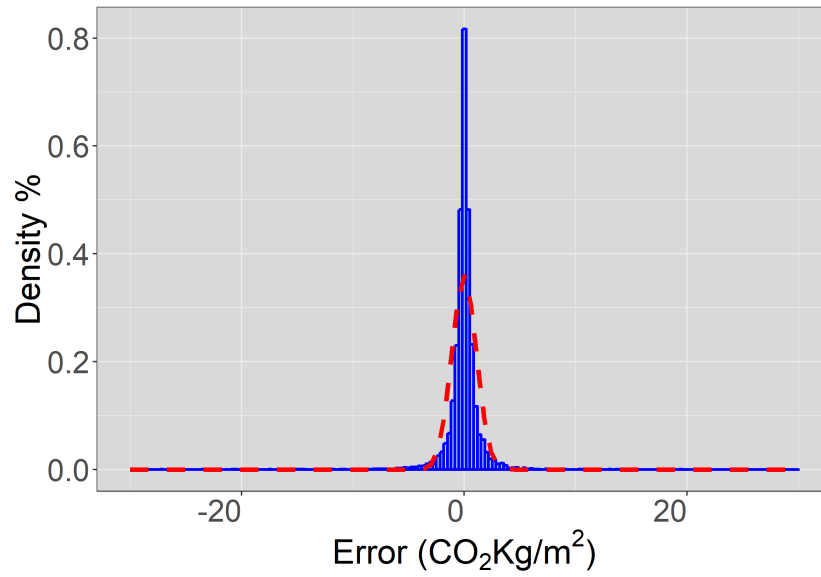
To demonstrate the performance of the model with dropping less important features (CI_{SW} , CI_{NW} , CI_{NE} , I_{NW} , I_{NE} , CI_{SE} , CI_{E} , I_{SE} , I_{SW} , CI_{W}), a model was trained using 30,000 records and tested over 8,000 samples. Both train and test sets were randomly selected among available 80,000 building records. The results are demonstrated as a plot of predicted energy performances versus real simulation values and distribution of error between simulated-predicted pairs in Figure 9.

4.4. The Case Study

The energy performance prediction model development aimed at supporting the optimisation of building energy retrofits. In this application, the optimisation algorithm generates variations of the target building to be evaluated for their energy performance. Here, the developed model is validated by estimating retrofit versions of a building that are never seen by



(a)



(b)

Figure 9: (a) Actual and predicted building emission rates and (b) error histogram for 8,000 test cases with a model trained over 30,000 samples.

the model. Therefore, the goal of this analysis is to further investigate the stability and generalization of machine learning (i.e. the training and testing the model is not biased by the utilised dataset).

To evaluate the efficiency of the developed model on the prediction of variations of a building, a single floor from a non-domestic building located in Glasgow city was selected. The details and data for the eighth floor of Graham Hills building were obtained from the Estates department of the University of Strathclyde. The building energy model was translated from IES Workstation to DesignBuilder before being extended to introduce the zoning shown in Figure 10.

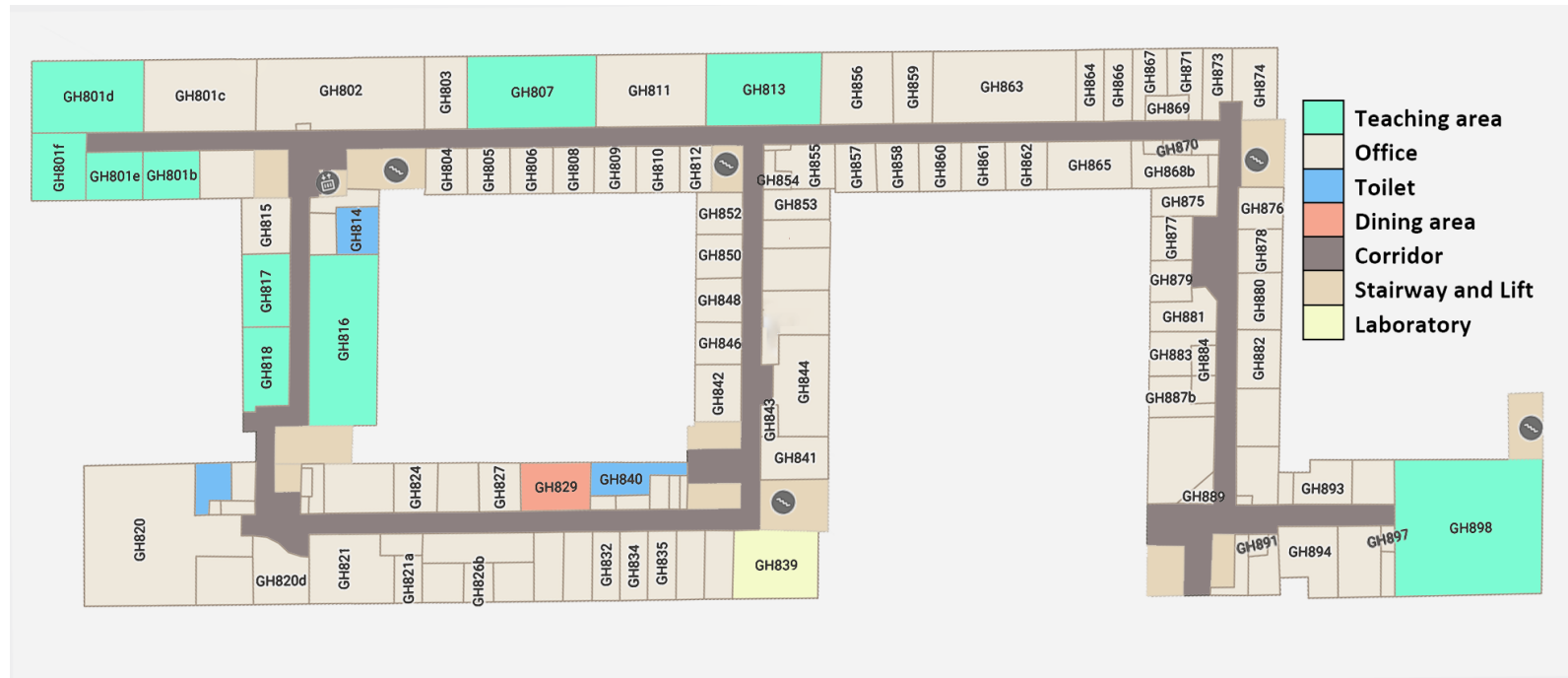


Figure 10: Case study floor plan showing the zones with activities

All zones are naturally ventilated except toilets which have local extract fans. Although the eighth floor exclusively uses low-temperature hot water boilers with wet radiators, other levels share several different systems, some of which including cooling and mechanical ventilation. The DHW system does not share heat generation with the wet radiator systems.

Considering the recommendations introduced in Table 1 and employing the GA, the studied building was mutated into 3,000 distinct retrofitted versions. All generated building models were assessed by the software and received the emission ratings, which vary from 51 to 80 $kgCO_2/(m^2 \times year)$. The process of evaluating all generated data using the SBEM tool took almost three days. In order to take full advantage of the ML model for this case study, the model was fitted using the whole training set. Then the trained model was used to predict the energy performance of the mutated records. Testing all samples took only 0.22 seconds while having RSME, MAE and R^2 of 1.02, 0.47 and 0.98, respectively.

Figure 11 demonstrates the performance of the predictions as the error histogram.

The obtained RMSE is equal to 1.7% of the average BER value, which is 58.3 $kgCO_2/m^2 \times year$. It can also be seen in Figure 11 that the majority of the residuals are cumulated around zero, proving the success of training a model with high accuracy and generality.

5. Conclusion

This research fills into the theoretical and practical gap regarding the insufficiency of existing multi-objective optimisation approaches in supporting deep energy retrofit of non-domestic buildings due to their complexity and scales which leads to high computational and labour cost. In accordance with this aim, it develops an energy performance prediction model for non-domestic buildings with the aid of machine learning to provide a rapid calculation engine for assisting energy retrofit decision-making. The developed model allows bypassing time-consuming energy simulations in non-domestic building energy retrofit optimisation to accelerate the process. This method makes use of AI in retrofit decision-making computationally feasible for complex buildings and on a large scale. The work contributes to carbon footprint reduction by decreasing the energy consumption, where both community and stockholders will be the beneficiaries of the provided service. The main

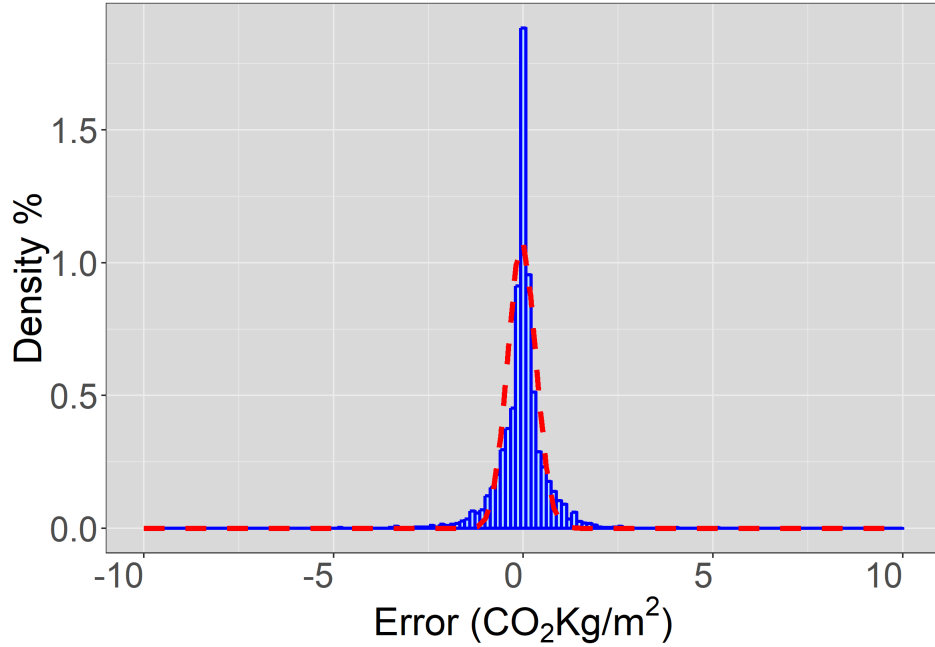


Figure 11: Error histogram of predicting 3,000 retrofit variations of the Graham Hill building

advantages of this method over utilisation of simulation software include a significant reduction in the time complexity of energy performance calculation, the potential to investigate a more comprehensive range of retrofit technologies, and consequently the ability to perform a proper inspection of retrofit exemptions denoted by the policies.

Although the development of data-driven models requires a considerable amount of historical data and precise optimisation to achieve high accuracy in prediction of building energy performance, the speed of calculations is not comparable with energy simulation software. The most time-consuming procedure in ML model development is feature engineering which can be considered as software development stage for energy simulation. Once the model is trained, it can assess millions of building records in a second. As the results indicated three thousands variations of a case study building which SBEM simulation took three days to calculate their emission rates, were assessed by the ML model in less than 0.22 seconds while providing a high accuracy as 1.7% of the average BER value. Meanwhile, the model can be

updated with new data very fast to improve accuracy further. It should be noted again that the ML model is not proposed to ultimately replace the engineering methods, but to promote the decision-making process where massive calculations are required. In practice, when a retrofit configuration or building design is selected with the support of a data-driven model, it should be evaluated by the base engineering method for the detail energy simulation (e.g. SBEM in this study).

In the proposed method, the essential characteristics of the non-domestic building affecting the energy performance were identified and formulated as a set of numerical features. This research investigated the retrofit recommendations before extracting features for machine learning modelling and explained how these technologies are affecting the model in prediction of building energy performance. Real-world data was then processed and translated into the described feature set for performing statistical modelling. The data was further improved by mutating the records and evaluating them using energy simulation software. The gradient boosted regression tree model was then trained and tuned and tested using cross-validation. The effectiveness of feature extraction and engineering was also evaluated by employing sensitivity analysis.

The proposed approach was evaluated using a functional non-domestic building as a case study and testing the accuracy of the machine learning model over the retrofit (mutated) versions which was simulated to get the emission rates. The precision of the developed model was validated through comparisons with the simulations. Development of an accurate model for estimation of the energy performance with a quick and robust process lays the groundwork for more informed and prolific consideration for deep energy retrofit decision-making.

The proposed model solves two significant problems with the existing tools which leverage SBEM, static retrofit package definitions and blindness to indirect service dependence. The former is a limitation on the size of the solution space that can be exhaustively explored. The latter are scenarios where retrofitting one building service increases demand on another to the extent that the energy savings for the isolated service demand are negated by the increased demand on the dependent service. The typical example of SBEM is the relationship between Tungsten or Halogen lighting and direct or storage electric heating where heating is the dependent service. As far as SBEM is concerned, lighting fixtures both contribute significantly to space heating and their contributions when considered with also meeting the

lighting demand, are more efficient than the electric heating system in tempering the space. The existing model, which works based on an exhaustive search of the solution space using linear estimation, treats the retrofit as an antagonistic matter. As such, it will not serve as a component of a candidate retrofit strategy. The proposed model solves and can leverage the underpinning diminishment of energy savings. It solves the problem by producing an estimate that is sufficiently accurate to be considered as a proxy for a nonlinear estimation of each retrofit scenario meaning a retrofit which might be antagonistic in isolation is considered in compound scenarios. The model, being many orders of magnitude faster than existing models, enables variable calibration of individual and grouped retrofit technologies. This can be leveraged during MOO with the nonlinear relationship between multiple objectives to find the Pareto fronts that approach losses to meet another objective.

In the previous section, it was discussed how the proposed model expands the solution space of the existing model. The solution space inherently changes every time what is known about the building or retrofits change. Therefore, it is desirable to calibrate the input parameters when a significant input is modified. For example, the model may suggest technology at an estimated price. The proposed model enables the introduction of tender specifications to MOO, offer a means of optimising the bid for the tendering party or enable the owner to hold tender-vs-tender MOO. Generally, the performance of the proposed model should facilitate consideration for occupant wellbeing benefits or other less conventional objectives.

The framework for feature extraction and analysis presented in this study can be adopted in any region with diverse climates. Besides, the SBEM which the proposed model estimate is a modified version of the Dutch methodology NEN 2916:1998 (Energy Performance of Non-Residential Buildings), whose outputs are energy in/out and carbon emissions, is constrained in energy performance estimation only by the climate data availability. The outputs have global relevance which suggests the model could be appropriate for analysis building stock that is not limited by any host country.

The results in this study are also validated by their accuracy in relation to the equivalent SBEM model's accredited energy performance. SBEM is the UK's model for fulfilling European Building Performance Directive Article 3 requirements for non-domestic buildings, and its results are approved for demonstrating compliance with Article 4 and demonstrating

compliance under Article 6 and 7. Additionally, SBEM's results are valid for demonstrating compliance with England and Wales' non-domestic minimum energy efficiency standards and energy savings opportunity scheme reporting, Scotland's Section 63 and Republic of Ireland's Non-domestic Energy Assessment Procedure.

This research highlights the significance of feature extraction and engineering in the estimation of energy indices applying the built-in mechanisms of gradient boosted regression trees. There have been several assumptions and limitations in creating building models and simulations that affect energy predictions. Accordingly, this study demonstrated that in the development of energy modelling for retrofit decision-making or other applications, such as building management, achieving a high accuracy is not the only concern. It is crucial to consider all the variations and reflects the corresponding impacts in the model behaviour. For example, a proper energy model for supporting building energy management system should consider occupant behaviour, as it can vary by different situations and cause model failure. Whilst the model can accurately perform in the normal condition due to insufficient data collection. The presented results also revealed the potential of machine learning-based modelling for removing irrelevant variables without influencing accuracy.

The research emphasises the capacity of machine learning methods in the built environment where calculation and simulation of energy indicators applying engineering methods sometimes become cumbersome. Thus far, it has been estimated that only 4% of data captured in the industrial environments are being employed with a meaningful and significant outcome. This is why Industry 4.0 has put more emphasis on the utilisation of technologies that could take advantage of the ever-growing data [87].

A technological innovation's short-term value can be measured by its ability to improve human capacity for an action in one or more metrics higher than its negation on other parameters. This paper presents an approach which on several metrics significantly improved researchers' capacity to exploit artificial intelligence optimisation effectively. It will facilitate significant improvements in model design and implementation. Reducing the duration of the process is essentially increasing the frequency of model calibrations which means there is less opportunity cost per calibration. It also increases the capacity to refine model creation processing for progressive interactive training models to reduce reliance hierarchical reinforcement learning.

As policy tightens on inefficient energy consumption and the understanding of the limitations of building energy management led retrofit decision-making, the necessity for more efficient and flexible models increases. Research over the last few years has been giving greater credence to designing buildings with consideration for medium-term climate change and any number of occupant presence or behaviour uncertainties. Every extension to the potential configurations exponentially inflates the problem space while likely reducing the conventional options solution space. Furthermore, these climate and utilisation properties are internal to building energy management. However, design and retrofit analysis is increasingly considering external and more challenging to integrate properties. The procedure introduced in this study demonstrated that algorithmic decision-making capabilities were not nearing their limit and laid a foundation for more complex machine learning frameworks. This study has significant implications on the industry in terms of supporting building design and construction practices as well as the energy policy-making. The building industry will also benefit from adopting the approaches demonstrated in this work to evaluate the effectiveness of the retrofitting methods in use and for the leveraging the performance of the energy modelling software applications.

Acknowledgement

The research presented in this paper was co-funded by The Data Lab (Edinburgh, UK) and arbnco Ltd (Glasgow, UK), through DataLab SFC Earmarked Grant Agreement: PO DL 00033. This work would also not be feasible without the generous PhD funding for the first author, which was co-funded by the Engineering The Future scheme from University of University of Strathclyde and the Industry Funded Studentship Agreement with arbnco Ltd (Studentship Agreement Number: S170392-101).

References

- [1] S. Kelly, D. Crawford-Brown, M. G. Pollitt, Building performance evaluation and certification in the UK: Is SAP fit for purpose?, *Renewable and Sustainable Energy Reviews* 16 (9) (2012) 6861–6878. doi:<http://dx.doi.org/10.1016/j.rser.2012.07.018>.

- [2] T. Hong, M. A. Piette, Y. Chen, S. H. Lee, S. C. Taylor-Lange, R. Zhang, K. Sun, P. Price, Commercial Building Energy Saver: An energy retrofit analysis toolkit, *Applied Energy* 159 (2015) 298–309. doi:10.1016/j.apenergy.2015.09.002.
- [3] T. Ibn-Mohammed, R. Greenough, S. Taylor, L. Ozawa-Meida, A. Acquaye, Integrating economic considerations with operational and embodied emissions into a decision support system for the optimal ranking of building retrofit options, *Building and Environment* 72 (2014) 82–101. doi:10.1016/j.buildenv.2013.10.018.
- [4] J. Ling-Chin, W. Taylor, P. Davidson, D. Reay, W. I. Nazi, S. Tassou, A. P. Roskilly, UK building thermal performance from industrial and governmental perspectives, *Applied Energy* 237 (2019) 270–282. doi:10.1016/j.apenergy.2018.12.077.
- [5] F. Ascione, N. Bianco, C. De Stasio, G. M. Mauro, G. P. Vanoli, A new methodology for cost-optimal analysis by means of the multi-objective optimization of building energy performance, *Energy and Buildings* 88 (2015) 78–90. doi:10.1016/j.enbuild.2014.11.058.
- [6] S. Oliver, S. Seyedzadeh, F. Pour Rahimian, Using real occupancy in retrofit decision-making: reducing the performance gap in low utilisation higher education buildings, in: 36th CIB W78 2019 Conference, 2019. URL https://strathprints.strath.ac.uk/69619/1/Oliver_etal_CIB2019_Using_real_occupancy_in_retrofit_decision_making.pdf, Accessed 30th June 2020
- [7] UK Department of Energy & Climate, Private Rented Sector Minimum Energy Efficiency Standard Regulations (Non- Domestic, England and Wales), Tech. Rep. July, Department of Energy and Climate Change (DECC)., London (2014). URL https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/401378/Non_Dom_PRS_Energy_Efficiency_Regulations_-_Gov_Response_FINAL_1_1_04_02_15_.pdf, Accessed 30th June 2020
- [8] S. Bu, G. Shen, C. J. Anumba, A. K. Wong, X. Liang, Literature review of green retrofit design for commercial buildings with BIM implication,

- Smart and Sustainable Built Environment 4 (2) (2015) 188–214. doi:10.1108/SASBE-08-2014-0043.
- [9] P. Pilechiha, M. Mahdaveinejad, F. Pour Rahimian, P. Carnemolla, S. Seyedzadeh, Multi-objective optimisation framework for designing office windows: quality of view, daylight and energy efficiency, *Applied Energy* 261 (2020) 114356. doi:10.1016/j.apenergy.2019.114356.
 - [10] D. B. Crawley, L. K. Lawrie, F. C. Winkelmann, W. F. Buhl, Y. J. Huang, C. O. Pedersen, R. K. Strand, R. J. Liesen, D. E. Fisher, M. J. Witte, J. Glazer, EnergyPlus: Creating a new-generation building energy simulation program, *Energy and Buildings* 33 (4) (2001) 319–331. doi:[https://doi.org/10.1016/S0378-7788\(00\)00114-6](https://doi.org/10.1016/S0378-7788(00)00114-6).
 - [11] TRNSYS17, A Transient Systems Simulation Program (2015).
URL <http://sel.me.wisc.edu/trnsys/>, Accessed 30th June 2020
 - [12] The Energy Systems Research Unit (ESRU), ESP-r (2011).
URL <http://www.esru.strath.ac.uk/Programs/ESP-r.htm>, Accessed 30th June 2020
 - [13] C. Pountney, Better carbon saving: using a genetic algorithm to optimise building carbon reductions, in: *Proceedings of the 2012 Building Simulation and Optimization Conference*, no. September, 2012, pp. 165–172.
URL <http://www.ibpsa.org/proceedings/BS02012/3A1.pdf>, Accessed 30th June 2020
 - [14] I. García Kerdan, R. Raslan, P. Ruyssevelt, An exergy-based multi-objective optimisation model for energy retrofit strategies in non-domestic buildings, *Energy* 117 (2016) 506–522. doi:10.1016/j.energy.2016.06.041.
 - [15] H. Adan, F. Fuerst, Modelling energy retrofit investments in the UK housing market: A microeconomic approach, *Smart and Sustainable Built Environment* 4 (3) (2015) 251–267. doi:10.1108/SASBE-03-2013-0016.
 - [16] W. Yu, B. Li, H. Jia, M. Zhang, D. Wang, Application of multi-objective genetic algorithm to optimize energy efficiency and thermal comfort in

- building design, *Energy and Buildings* 88 (2015) 135–143. doi:10.1016/j.enbuild.2014.11.063.
- [17] F. Ascione, N. Bianco, G. M. Mauro, G. P. Vanoli, A new comprehensive framework for the multi-objective optimization of building energy design: Harlequin, *Applied Energy* 241 (2019) 331–361. doi:10.1016/j.apenergy.2019.03.028.
 - [18] M. Ferrara, A. Rolfo, F. Prunotto, E. Fabrizio, EDeSSOpt Energy Demand and Supply Simultaneous Optimization for cost-optimized design: Application to a multi-family building, *Applied Energy* 236 (2019) 1231–1248. doi:10.1016/j.apenergy.2018.12.043.
 - [19] HM Government, Approved Document L2B: conservation of fuel and power in existing buildings other than dwellings, 2010 edition (incorporating 2010, 2011, 2013 and 2016 amendments), Tech. rep., Ministry of Housing, Communities & Local Government (2016).
URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/540329/BR_PDF_AD_L2B_2013_with_2016_amendments.pdf, Accessed 30th June 2020
 - [20] S. Wang, C. Yan, F. Xiao, Quantitative energy performance assessment methods for existing buildings, *Energy and Buildings* 55 (2012) 873–888. doi:http://dx.doi.org/10.1016/j.enbuild.2012.08.037.
 - [21] S. M. Hong, G. Paterson, E. Burman, P. Steadman, D. Mumovic, A comparative study of benchmarking approaches for non-domestic buildings: Part 1 Top-down approach, *International Journal of Sustainable Built Environment* 2 (2) (2013) 119–130. doi:10.1016/j.ijlsbe.2014.04.001.
 - [22] E. Burman, S. M. Hong, G. Paterson, J. Kimpian, D. Mumovic, A comparative study of benchmarking approaches for non-domestic buildings: Part 2 Bottom-up approach, *International Journal of Sustainable Built Environment* 3 (2) (2014) 247–261. doi:10.1016/j.ijlsbe.2014.12.001.
 - [23] M. Ouf, M. H. Issa, Energy consumption analysis of school buildings in Manitoba, Canada, *International Journal of Sustainable Built*

- Environment 6 (2) (2017) 359–371. doi:<https://doi.org/10.1016/j.ijlsbe.2017.05.003>.
- [24] E. H. Borgstein, R. Lamberts, J. L. Hensen, Evaluating energy performance in non-domestic buildings: A review, *Energy and Buildings* 128 (2016) 734–755. doi:<https://doi.org/10.1016/j.enbuild.2016.07.018>.
 - [25] Coakley, P. Raftery, M. Keane, A review of methods to match building energy simulation models to measured data, *Renewable & Sustainable Energy Reviews* 37 (2014) 123–141. doi:<https://doi.org/10.1016/j.rser.2014.05.007>.
 - [26] S. Seyedzadeh, F. P. Rahimian, I. Glesk, M. Roper, Machine learning for estimation of building energy consumption and performance: a review, *Visualization in Engineering* 6 (1) (2018) 5. doi:[10.1186/s40327-018-0064-7](https://doi.org/10.1186/s40327-018-0064-7).
 - [27] J. S. Hygh, J. F. DeCarolis, D. B. Hill, S. R. Ranjithan, Multivariate regression as an energy assessment tool in early building design, *Building and Environment* 57 (2012) 165–175. doi:<https://doi.org/10.1016/j.buildenv.2012.04.021>.
 - [28] D. Ruch, L. Chen, J. S. Haberl, D. E. Claridge, A Change-Point Principal Component Analysis (CP/PCA) Method for Predicting Energy Usage in Commercial Buildings: The PCA Model, *Journal of Solar Energy Engineering* 115 (2) (1993) 77. doi:<https://doi.org/10.1115/1.2930035>.
 - [29] W. Meeusen, J. van Den Broeck, Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error, *International Economic Review* 18 (2) (1977) 435. doi:<https://doi.org/10.2307/2525757>.
 - [30] W.-S. Lee, K.-P. Lee, Benchmarking the performance of building energy management using data envelopment analysis, *Applied Thermal Engineering* 29 (16) (2009) 3269–3273. doi:<https://doi.org/10.1016/j.applthermaleng.2008.02.034>.
 - [31] S. H. Mousavi-Avval, S. Rafiee, A. Jafari, A. Mohammadi, Optimization of energy consumption for soybean production using Data Envelopment

- Analysis (DEA) approach, *Applied Energy* 88 (11) (2011) 3765–3772. doi:<https://doi.org/10.1016/j.apenergy.2011.04.021>.
- [32] S. Seyedzadeh, F. P. Rahimian, S. Oliver, I. Glesk, B. Kumar, Data Driven Model Improved by Multi-Objective Optimisation for Prediction of Building Energy Loads Teesside University, *Automation in Construction* 116 (2020) 103188. doi:<https://doi.org/10.1016/j.autcon.2020.103188>.
 - [33] H. X. Zhao, F. Magoulès, Feature selection for predicting building energy consumption based on statistical learning method, *Journal of Algorithms and Computational Technology* 6 (1) (2012) 59–77. doi:10.1260/1748-3018.6.1.59.
 - [34] M. Yalcintas, U. A. Ozturk, An energy benchmarking model based on artificial neural network method utilizing US Commercial Buildings Energy Consumption Survey (CBECS) database, *International Journal of Energy Research* 31 (4) (2007) 412–421. doi:10.1002/er.1232.
 - [35] M. Yalcintas, An energy benchmarking model based on artificial neural network method with a case example for tropical climates, *International Journal of Energy Research* 30 (14) (2006) 1158–1174. doi:10.1002/er.1212.
 - [36] S. L. Wong, K. K. W. Wan, T. N. T. Lam, Artificial neural networks for energy analysis of office buildings with daylighting, *Applied Energy* 87 (2) (2010) 551–557. doi:10.1016/j.apenergy.2009.06.028.
 - [37] M. Lundin, S. Andersson, R. Åstin, Development and validation of a method aimed at estimating building performance parameters, *Energy and Buildings* 36 (9) (2004) 905–914. doi:<https://doi.org/10.1016/j.enbuild.2004.02.005>.
 - [38] C. Buratti, M. Barbanera, D. Palladino, An original tool for checking energy performance and certification of buildings by means of Artificial Neural Networks, *Applied Energy* 120 (2014) 125–132. doi:10.1016/j.apenergy.2014.01.053.
 - [39] S.-M. M. Hong, G. Paterson, D. Mumovic, P. Steadman, Improved benchmarking comparability for energy consumption in schools,

- Building Research & Information 42 (1) (2014) 47–61. doi:<https://doi.org/10.1080/09613218.2013.814746>.
- [40] F. Khayatian, L. Sarto, G. Dall’O’, Application of neural networks for evaluating energy performance certificates of residential buildings, *Energy and Buildings* 125 (2016) 45–54. doi:10.1016/j.enbuild.2016.04.067.
 - [41] F. Ascione, N. Bianco, C. De Stasio, G. M. Mauro, G. P. Vanoli, Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach, *Energy* 118 (2017) 999–1017. doi:10.1016/j.energy.2016.10.126.
 - [42] M. Beccali, G. Ciulla, V. Lo Brano, A. Galatioto, M. Bonomolo, Artificial neural network decision support tool for assessment of the energy performance and the refurbishment actions for the non-residential building stock in Southern Italy, *Energy* 137 (2017) 1201–1218. doi:10.1016/j.energy.2017.05.200.
 - [43] Q. Li, P. Ren, Q. Meng, Prediction model of annual energy consumption of residential buildings, in: 2010 International Conference on Advances in Energy Engineering, ICAEE 2010, IEEE, 2010, pp. 223–226. doi:10.1109/ICAEE.2010.5557576.
 - [44] D. M. Solomon, R. L. Winter, A. G. Boulanger, R. N. Anderson, L. L. Wu, Forecasting energy demand in large commercial buildings using support vector machine regression, Tech. rep. (2011).
URL <http://academiccommons.columbia.edu/catalog/ac:143154>, Accessed 30th June 2020
 - [45] A. Tsanas, A. Xifara, Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools, *Energy and Buildings* 49 (2012) 560–567. doi:10.1016/j.enbuild.2012.03.003.
 - [46] S. Papadopoulos, E. Azar, W.-L. Woon, C. E. Kontokosta, Evaluation of tree-based ensemble learning algorithms for building energy performance estimation, *Journal of Building Performance Simulation* 1493 (2017) 1–11. doi:<https://doi.org/10.1080/19401493.2017.1354919>.

- [47] H. Deng, D. Fannon, M. J. Eckelman, Predictive modeling for US commercial building energy use: A comparison of existing statistical and machine learning algorithms using CBECS microdata, *Energy and Buildings* 163 (2018) 34–43. doi:<https://doi.org/10.1016/j.enbuild.2017.12.031>.
- [48] Z. Wang, Y. Wang, R. Zeng, R. S. Srinivasan, S. Ahrentzen, Random Forest based hourly building energy prediction, *Energy and Buildings* 171 (2018) 11–25. doi:<https://doi.org/10.1016/j.enbuild.2018.04.008>.
- [49] S. Seyedzadeh, F. Pour Rahimian, P. Rastogi, I. Glesk, Tuning machine learning models for prediction of building energy loads, *Sustainable Cities and Society* 47. doi:[10.1016/j.scs.2019.101484](https://doi.org/10.1016/j.scs.2019.101484).
- [50] S. Papadopoulos, E. Azar, W. L. Woon, C. E. Kontokosta, Evaluation of tree-based ensemble learning algorithms for building energy performance estimation, *Journal of Building Performance Simulation* 11 (3) (2018) 322–332. doi:[10.1080/19401493.2017.1354919](https://doi.org/10.1080/19401493.2017.1354919).
- [51] J. Field, J. Soper, P. Jones, W. Bordass, P. Grigg, Energy performance of occupied non-domestic buildings: Assessment by analysing end-use energy consumptions, *Building Services Engineering Research and Technology* 18 (1) (1997) 39–46. doi:<https://doi.org/10.1177/014362449701800106>.
- [52] R. Bull, N. Chang, P. Fleming, The use of building energy certificates to reduce energy consumption in European public buildings, *Energy and Buildings* 50 (2012) 103–110. doi:[10.1016/j.enbuild.2012.03.032](https://doi.org/10.1016/j.enbuild.2012.03.032).
- [53] Communities and Government Local: London, Performance Testing of Buildings BD 2535, Tech. rep., Department for Communities and Local Government, London, UK (2013).
URL <https://www.thenbs.com/PublicationIndex/documents/details?Pub=DCLG&DocID=286091>, Accessed 30th June 2020
- [54] C. Koo, T. Hong, M. Lee, H. Park, Development of a new energy efficiency rating system for existing residential buildings, *Energy Policy* 68 (2014) 218–231. doi:<https://doi.org/10.1016/j.enpol.2013.12.068>.

- [55] Building Research Establishment (BRE), Building Standards Division National Calculation Methodology (NCM) Modelling Guide for Non-Domestic Buildings in Scotland, Tech. rep. (2015).
URL <http://www.gov.scot/bsd>, Accessed 30th June 2020
- [56] E. Abel, A. Elmroth, Buildings and Energy - a systematic approach, Vol. 6, Formas, 2007.
URL <http://lup.lub.lu.se/record/8056305>, Accessed 30th June 2020
- [57] E. Asadi, M. G. Da Silva, C. H. Antunes, L. Dias, Multi-objective optimization for building retrofit strategies: A model and an application, *Energy and Buildings* 44 (1) (2012) 81–87. doi:[10.1016/j.enbuild.2011.10.016](https://doi.org/10.1016/j.enbuild.2011.10.016).
- [58] C. Diakaki, E. Grigoroudis, D. Kolokotsa, Towards a multi-objective optimization approach for improving energy efficiency in buildings, *Energy and Buildings* 40 (9) (2008) 1747–1754. doi:<https://doi.org/10.1016/j.enbuild.2008.03.002>.
- [59] J. S. Gero, N. D’Cruz, A. D. Radford, Energy in context: A multicriteria model for building design, *Building and Environment* 18 (3) (1983) 99–107. doi:[https://doi.org/10.1016/0360-1323\(83\)90001-X](https://doi.org/10.1016/0360-1323(83)90001-X).
- [60] S. Pohekar, M. Ramachandran, Application of multi-criteria decision making to sustainable energy planningA review, *Renewable and Sustainable Energy Reviews* 8 (4) (2004) 365–381. doi:<https://doi.org/10.1016/j.rser.2003.12.007>.
- [61] J. K. W. Wong, H. Li, Application of the analytic hierarchy process (AHP) in multi-criteria analysis of the selection of intelligent building systems, *Building and Environment* 43 (1) (2008) 108–125. doi:<https://doi.org/10.1016/j.buildenv.2006.11.019>.
- [62] D. Kolokotsa, C. Diakaki, E. Grigoroudis, G. Stavrakakis, K. Kalaitzakis, Decision support methodologies on the energy efficiency and energy management in buildings, *Advances in Building Energy Research* 3 (1) (2009) 121–146. doi:<https://doi.org/10.3763/aber.2009.0305>.

- [63] H. Ren, W. Gao, W. Zhou, Multi-criteria evaluation for the optimal adoption of distributed residential energy systems in Japan, *Energy Policy* 37 (12) (2009) 5484–5493. doi:<https://doi.org/10.1016/j.enpol.2009.08.014>.
- [64] M. Jaggs, J. Palmer, Energy performance indoor environmental quality retrofit - a European diagnosis and decision making method for building refurbishment, *Energy and Buildings* 31 (2) (2000) 97–101. doi:[https://doi.org/10.1016/S0378-7788\(99\)00023-7](https://doi.org/10.1016/S0378-7788(99)00023-7).
- [65] F. Flourentzou, C. A. Roulet, Elaboration of retrofit scenarios, *Energy and Buildings* 34 (2) (2002) 185–192. doi:[https://doi.org/10.1016/S0378-7788\(01\)00106-2](https://doi.org/10.1016/S0378-7788(01)00106-2).
- [66] A. Kaklauskas, E. K. Zavadskas, S. Raslanas, Multivariant design and multiple criteria analysis of building refurbishments, *Energy and Buildings* 37 (4) (2005) 361–372. doi:<https://doi.org/10.1016/j.enbuild.2004.07.005>.
- [67] E. Asadi, M. G. da Silva, C. H. Antunes, L. Dias, State of the Art on Retrofit Strategies Selection Using Multi-objective Optimization and Genetic Algorithms, in: *Nearly Zero Energy Building Refurbishment*, Springer, 2013, pp. 279–297. doi:https://doi.org/10.1007/978-1-4471-5523-2_11.
- [68] E. Asadi, M. G. D. Silva, C. H. Antunes, L. Dias, L. Glicksman, Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application, *Energy and Buildings* 81 (2014) 444–456. doi:[10.1016/j.enbuild.2014.06.009](https://doi.org/10.1016/j.enbuild.2014.06.009).
- [69] C. Diakaki, E. Grigoroudis, N. Kabelis, D. Kolokotsa, K. Kalaitzakis, G. Stavrakakis, A multi-objective decision model for the improvement of energy efficiency in buildings, *Energy* 35 (12) (2010) 5483–5496. doi:<https://doi.org/10.1016/j.energy.2010.05.012>.
- [70] E. Asadi, M. G. da Silva, C. H. Antunes, L. Dias, A multi-objective optimization model for building retrofit strategies using TRNSYS simulations, GenOpt and MATLAB, *Building and Environment* 56 (2012) 370–378. doi:[10.1016/j.buildenv.2012.04.005](https://doi.org/10.1016/j.buildenv.2012.04.005).

- [71] F. Ascione, N. Bianco, C. De Stasio, G. M. Mauro, G. P. Vanoli, Simulation-based model predictive control by the multi-objective optimization of building energy performance and thermal comfort, *Energy and Buildings* 111 (2016) 131–144. doi:10.1016/j.enbuild.2015.11.033.
- [72] S. Gou, V. M. Nik, J. L. Scartezzini, Q. Zhao, Z. Li, Passive design optimization of newly-built residential buildings in Shanghai for improving indoor thermal comfort while reducing building energy demand, *Energy and Buildings* 169 (2018) 484–506. doi:10.1016/j.enbuild.2017.09.095.
- [73] F. Bre, V. D. Fachinotti, A computational multi-objective optimization method to improve energy efficiency and thermal comfort in dwellings, *Energy and Buildings* 154 (2017) 283–294. doi:10.1016/j.enbuild.2017.08.002.
- [74] A. Jafari, V. Valentin, An optimization framework for building energy retrofits decision-making, *Building and Environment* 115 (2017) 118–129. doi:10.1016/j.buildenv.2017.01.020.
- [75] S. Carlucci, G. Cattarin, F. Causone, L. Pagliano, Multi-objective optimization of a nearly zero-energy building based on thermal and visual discomfort minimization using a non-dominated sorting genetic algorithm (NSGA-II), *Energy and Buildings* 104 (2015) 378–394. doi:10.1016/j.enbuild.2015.06.064.
- [76] F. Dell’Anna, M. Bottero, C. Becchio, S. P. Corgnati, G. Mondini, Designing a decision support system to evaluate the environmental and extra-economic performances of a nearly zero-energy building, *Smart and Sustainable Built Environment* doi:10.1108/SASBE-09-2019-0121.
- [77] J. H. Friedman, Stochastic gradient boosting, *Computational Statistics and Data Analysis* 38 (4) (2002) 367–378. doi:10.1016/S0167-9473(01)00065-2.
- [78] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, *Classification and regression trees*, Routledge, 2017. doi:10.1201/9781315139470.

- [79] S. Seyedzadeh, P. Rastogi, F. Pour Rahimian, S. Oliver, I. Glesk, B. Kumar, Multi-Objective Optimisation for Tuning Building Heating and Cooling Loads Forecasting Models, in: 36th CIB W78 2019 Conference, Newcastle, 2019.
URL https://strathprints.strath.ac.uk/69599/1/Seyedzadeh_etal_CIB2019_Multi_objective_optimisation_for_tuning_building_heating_and_cooling_loads.pdf, Accessed 30th June 2020
- [80] K. Deb, Multi-objective optimization, in: Search Methodologies, Springer, Boston, 2013, pp. 403–449. doi:https://doi.org/10.1007/978-1-4614-6940-7_15.
- [81] F. Hutter, H. H. Hoos, K. Leyton-Brown, Sequential model-based optimization for general algorithm configuration, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 6683 LNCS, Springer, 2011, pp. 507–523. doi:[10.1007/978-3-642-25566-3_40](https://doi.org/10.1007/978-3-642-25566-3_40).
- [82] A. Ng, Machine Learning and AI via Brain Simulations (2013).
URL <http://www.youtube.com/watch?v=n1ViNeWhC24%5Cnpapers3://publication/uuid/FD4F0FE0-BB23-431D-8840-6808135CC089>, Accessed 30th June 2020
- [83] Arbnco, arbn consult.
URL <https://arbnco.co.uk/arbn-consult/>, Accessed 30th June 2020
- [84] V. F. Rodriguez-Galiano, J. A. Luque-Espinar, M. Chica-Olmo, M. P. Mendes, Feature selection approaches for predictive modelling of groundwater nitrate pollution: An evaluation of filters, embedded and wrapper methods, Science of the Total Environment 624 (2018) 661–672. doi:[10.1016/j.scitotenv.2017.12.152](https://doi.org/10.1016/j.scitotenv.2017.12.152).
- [85] J. Benesty, J. Chen, Y. Huang, I. Cohen, Pearson Correlation Coefficient, in: Noise reduction in speech processing, Springer, 2009, pp. 1–4. doi:[10.1007/978-3-642-00296-0_5](https://doi.org/10.1007/978-3-642-00296-0_5).
- [86] L. Auret, C. Aldrich, Empirical comparison of tree ensemble variable importance measures, Chemometrics and Intelligent Laboratory

Systems 105 (2) (2011) 157–170. doi:10.1016/j.chemolab.2010.12.004.

- [87] C. Newman, D. Edwards, I. Martek, J. Lai, W. D. Thwala, I. Rillie, Industry 4.0 deployment in the construction industry: a bibliometric literature review and UK-based case study, Smart and Sustainable Built Environmentdoi:10.1108/SASBE-02-2020-0016.